RESEARCH ON SPOKEN LANGUAGE PROCESSING

Progress Report No. 25
(2001-2002)

David B. Pisoni, Ph.D.
Principal Investigator

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405-1301

Research Supported by:

Department of Health and Human Services
U.S. Public Health Service

National Institutes of Health
Research Grant No. DC-00111

and

National Institutes of Health
Training Grant No. DC-00012

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

Table of Contents

INTRODUCTION

This is the twenty-fifth annual progress report summarizing research activities on speech perception and spoken language processing carried out in the Speech Research Laboratory, Department of Psychology, Indiana University in Bloomington. As with previous reports, our main goal has been to summarize our accomplishments over the past year and make them readily available to granting agencies, sponsors and interested colleagues in the field. Some of the papers contained in this report are extended manuscripts that have been prepared for formal publication as journal articles or book chapters. Other papers are simply short reports of research presented at professional meetings during the past year or brief summaries of "on-going" research projects in the laboratory. From time to time, we also have included new information on instrumentation and software developments when we think this information would be of interest or help to others. We have found the sharing of this information to be very useful in facilitating research.

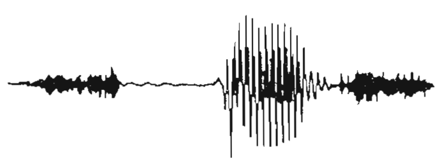
We are distributing progress reports of our research activities because of the ever increasing lag in journal publications and the resulting delay in the dissemination of new information and research findings in the field of spoken language processing. We are, of course, very interested in following the work of other colleagues who are carrying out research on speech perception and spoken language processing and we would be grateful if you and your colleagues would send us copies of any recent reprints, preprints and progress reports as they become available so that we can keep up with your latest findings. Please address all correspondence to:

Professor David B. Pisoni
Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405-1301
United States of America

Telephone: (812) 855-1155, 855-1768
Facsimile: (812) 855-1300
E-mail: pisoni@indiana.edu
Web: http://www.indiana.edu/~srlweb

Copies of this progress report are being sent primarily to libraries and specific research institutions rather than individual scientists. Because of the rising costs of publication and printing, it is not possible to provide multiple copies of this report to people at the same institution or issue copies to individuals. We are eager to enter into exchange agreements with other institutions for their reports and publications. Please write to the above address for further information.

The information contained in this progress report is freely available to the public and is not restricted in any way. The views expressed in these research reports are those of the individual authors and do not reflect the opinions of the granting agencies or sponsors of the specific research.



SPEECH – THE FINAL FRONTIER

SPEECH RESEARCH LABORATORY
FACULTY, STAFF, AND TECHNICAL PERSONNEL

(January 1, 2001–December 31, 2002)

RESEARCH PERSONNEL

David B. Pisoni, Ph.D. .... Chancellors' Professor of Psychology and Cognitive Science[1,2]

Karen I. Kirk, Ph.D. .... Associate Professor of Otolaryngology–Head and Neck Surgery[3,4]
Mario A. Svirsky, Ph.D. .... Associate Professor of Otolaryngology–Head and Neck Surgery[3,5]
Steven B. Chin, Ph.D. .... Assistant Scientist in Otolaryngology–Head and Neck Surgery[3]
Derek Houston, Ph.D. .... Assistant Professor of Otolaryngology–Head and Neck Surgery[3]

Woo Sui Teoh, M.D. .... NIH Postdoctoral Trainee[3]
Tonya Bergeson, Ph.D. .... NIH Postdoctoral Trainee[3]
Allyson K. Carter, Ph.D. .... NIH Postdoctoral Trainee

Rose Burkholder, B.S. .... NIH Predoctoral Trainee
Miranda Cleary, B.A. .... NIH Predoctoral Trainee
Cynthia G. Clopper, B.A. .... NIH Predoctoral Trainee
Brianna Conrey, B.A. .... NSF Predoctoral Trainee
Caitlin M. Dillon, B.A. .... NIH Predoctoral Trainee
Lorin Lachs, B.A. .... NIH Predoctoral Trainee
Noah Silbert, B.A. .... Predoctoral Trainee

TECHNICAL PERSONNEL

Luis R. Hernández, B.A. .... Research Associate in Psychology
Darla J. Sallee .... Administrative Assistant
Kuai Hinojosa .... Applications Programmer

Elena Breiter .... Undergraduate Research Assistant
Jeff Karpicke .... Undergraduate Research Assistant
Jeff Reynolds .... Undergraduate Research Assistant
Adam Tierney .... Undergraduate Research Assistant

---

[1] Also Adjunct Professor of Linguistics, Indiana University, Bloomington, IN.
[2] Also Adjunct Professor of Otolaryngology–Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, IN.
[3] Department of Otolaryngology–Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, IN.
[4] Also Adjunct Associate Professor of Speech and Hearing Sciences, Indiana University, Bloomington, IN.
[5] Also Adjunct Associate Professor of Electrical Engineering, Purdue School of Engineering and Technology, Indianapolis, IN.

E-Mail Addresses

Tonya Bergeson .... tbergeso@iupui.edu
Rose Burkholder .... rburkhol@indiana.edu
Allyson K. Carter .... allcarte@indiana.edu
Steven B. Chin .... schin@iupui.edu
Miranda Cleary .... micleary@indiana.edu
Cynthia G. Clopper .... cclopper@indiana.edu
Caitlin M. Dillon .... cmdillon@indiana.edu
Luis R. Hernández .... hernande@indiana.edu
Derek Houston .... dmhousto@indiana.edu
Jeff Karpicke .... karpicke@indiana.edu
Karen I. Kirk .... kkirk@iupui.edu
Lorin Lachs .... llachs@indiana.edu
David B. Pisoni .... pisoni@indiana.edu
Darla J. Sallee .... dsallee@indiana.edu
Mario A. Svirsky .... msvirsky@iupui.edu
Woo Sui Teoh .... stech@iupui.edu

Some New Findings on Learning, Memory and Cognitive Processes in Deaf Children Following Cochlear Implantation[1]

David B. Pisoni and Miranda Cleary

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

[1] This work was supported by NIH-NIDCD Training Grant T32 DC00012 and NIH-NIDCD Research Grants R01 DC00111, NIH NIDCD R01 DC00064 to Indiana University and NIH Research Grant R01 DC03100 to Central Institute for the Deaf. We would like to thank Tonya Bergeson, Steven Chin and Karen Iler Kirk for comments on an earlier version of this paper. This is a draft of a chapter to appear in F.G. Zeng, A.N. Popper & R.R. Fay (Eds.), *Handbook of Auditory Research: Auditory Prostheses*.

Some New Findings on Learning, Memory and Cognitive Processes in Deaf Children Following Cochlear Implantation

Abstract. The present chapter reports new findings on learning, memory and cognitive processes in deaf children following cochlear implantation that attempt to account for the large individual differences in acquisition of aural/oral language skills in this clinical population. The role of known contributing demographic factors is briefly reviewed and the need for further process measures of performance is presented. Perception, attention, memory, and learning in normal language development are discussed. It is suggested that these specific cognitive processes should be studied in greater detail in this clinical population in order to explain the enormous variability in spoken language acquisition. Results from two new data sets are presented and discussed. In the first study, results from an investigation of the "Stars," deaf children who do exceptionally well with their implant is presented and discussed. The exceptionally good implant users differed from the poorer implant users in several ways related to their ability to rapidly encode sound patterns into phonological representations in working memory and coordinate receptive verbal abilities with expressive language skills. Using supporting literature on typically-developing children, it is argued that a cognitive processing factor having to do with fast and efficient encoding, maintenance, and retrieval of phonological information in short-term is what enables the better-performing implant users to take advantage of this coordinated linguistic system. The second set of studies involved implanted children who showed a more typical range of abilities. Results from a variety of new measures of information processing performance are reported along with data from age-matched normal-hearing children and normal-hearing adults. Within the group of deaf children who use cochlear implants, speaking rate, a measure shown in other populations to correlate well with an individual's rehearsal speed for items in immediate memory was found to be strongly correlated with measures of a child's short-term memory capacity as well as his auditory-only spoken word recognition performance. Finally, additional evidence is presented suggesting atypical short-term memory spans for auditory as well as visual sequences in at least some deaf children with cochlear implants. Data from a novel sequence-learning task revealed that deaf children with cochlear implants also show less than expected benefit from the repetition of familiarized sequences. These new findings on learning, memory and cognitive processes suggest that variation in children's success with cochlear implants may reflect differences in the operation of basic information processing skills used in a wide range of language processing tasks that draw on verbal rehearsal processes used in working memory.

Introduction

Cochlear implants work reasonably well in many profoundly deaf adults and children. For these patients, a cochlear implant is a form of intervention, an alternative way of providing access to sound via electrical stimulation of the auditory system. For the post-lingually deafened adult, a cochlear implant serves primarily as a sensory aid to restore lost hearing and regain contact with the world of sound as they knew it before the onset of deafness. In contrast, for the prelingually deaf child, the electrical stimulation provided by a cochlear implant represents the introduction of a new sensory modality and an additional way to acquire knowledge about sound, sound sources and the correlations between objects and events in the environment. Perhaps the most important benefit of a cochlear implant, however, is that it provides the prelingually deaf child with access to information about speech and spoken language. Because speech is a multi-modal event, the electrical stimulation provided by the cochlear implant also provides the child with a rich source of new information about the cross-modal relations between the auditory and optical correlates of speech that reflect the common underlying articulatory gestures of the talker. Finally, a cochlear implant provides the deaf child with auditory feedback about the consequences of his own vocal articulation in speech production that affects the development of speech and language acquisition after implantation.

Despite the success of cochlear implants in many deaf patients, enormous individual differences have been reported in adults and children on a wide range of outcome measures. This finding is observed in all research centers around the world. Some patients do extremely well with their cochlear implants while others derive only minimal benefits after receiving their implants. Understanding the reasons for the variability in outcomes and the large individual differences following cochlear implantation is one of the most important and challenging research problems in the field today. It is not immediately obvious why some patients do well while others struggle and achieve only small benefits after receiving a cochlear implant. Many factors may be responsible for these differences, and numerous complex interactions among these factors should be explored.

Our initial interest in studying individual differences in children following cochlear implantation came from several reports in the literature demonstrating that a small number of deaf children displayed exceptionally good performance with their cochlear implants. They appeared to acquire spoken language quickly and easily and seemed to be on a developmental trajectory that paralleled children with normal hearing. These children are often called "Stars," and until recently their exceptionally good performance appeared to be an anomaly to many clinicians and researchers.

The finding that some deaf children with cochlear implants display exceptionally good performance can be taken, at first glance, as an "existence proof" for the efficacy of cochlear implants: cochlear implants work well with some children and they facilitate the processes of speech perception and language development. The major problem, however, is that cochlear implants do not work well with all children, and some children derive only minimal benefits from their implants. Why does this occur? What sensory, perceptual, cognitive and environmental factors are responsible for the differences in performance among deaf children with cochlear implants? These are the major questions that we have focused our research on over the last few years.

Our theoretical motivation for studying individual differences following cochlear implantation is based on an extensive body of research in the field of Cognitive Psychology over the last twenty-five years on "expertise" and "expert systems" theory (see Ericsson & Pennington, 1993). Many important new insights have come from studying expert chess players, radiologists and other people who have highly developed skills in specific knowledge domains like computer programming, spectrogram reading, and even chicken-sexing (see Biederman & Shiffrar, 1987).

The rationale underlying our approach is quite straightforward. If we can learn more about the exceptionally good users of cochlear implants and the reasons why they do so well, perhaps we can use this information to develop new intervention techniques with children who are not benefiting from their implants. Knowledge and understanding of the exceptionally good users, the "Stars," might also be useful for developing new pre-implant predictors of performance, modifying current criteria for candidacy and creating better methods of assessing performance and measuring outcome and benefit after implantation. Thus, there are many important clinical benefits that might result from research on individual differences in these particular children.

The variability in outcomes and the large individual differences following cochlear implantation also have implications for several important theoretical issues dealing with neural plasticity and development. Deaf children who receive cochlear implants have been deprived of sound input for some length of time after birth and their nervous systems have continued to develop in the absence of "normal" sensory stimulation during the critical period for language learning. These children represent a unique clinical population to study because they can provide important new information about the effects of early auditory deprivation on cognitive and linguistic development. What happens to the nervous system and brain of these children as a result of deafness and lack of auditory stimulation over this period of time? Can their atypical pattern of development be modified or reversed after the introduction of sound?

At the present, we know that a small number of demographic factors are strongly associated with a variety of speech and language development outcome measures in these children. However, the investigation of "higher-level" perceptual, cognitive and linguistic factors has not received very much attention until recently. One reason for our lack of knowledge about cognitive processes is that most of the clinical research on cochlear implants over the last 10-15 years has been carried out by audiologists and speech-language pathologists who have been concerned primarily with questions of device "efficacy" and assessment of outcome. Their primary interests have been focused on demonstrating that cochlear implants work and provide benefit to deaf patients. Historically, these researchers have had little if any interest in variation and individual differences in performance. Research on treatment efficacy requires well-defined assessment measures of outcome performance that are familiar to surgeons and clinicians who work with deaf patients. In contrast, research on variability and individual differences in performance deals with a somewhat different problem, namely the clinical "effectiveness" of cochlear implants, that is, explaining why cochlear implants do not work well in all patients who receive them.

Interest in individual differences following cochlear implantation has also become a high priority of the federal government, which funds basic and clinical research on hearing and deafness. In 1995, the National Institute of Health published a "Consensus Statement on Cochlear Implants in Adults and Children" to provide clinicians with an up-to-date summary of the benefits and limitations of cochlear implants (NIH, 1995). The NIH panel concluded that while cochlear implants improve communication abilities in most postlingually deafened adults with severe to profound hearing loss, the outcomes of implantation are much more variable in children, especially prelingually deafened children. Among other findings related to the efficacy and effectiveness of cochlear implants, the panel focused on the wide variation in outcome measures in implant users and recommended that additional basic and clinical research be carried out on individual differences in both adults and children. The panel also suggested that new methods and tools should be developed to study how cochlear implants activate the central auditory system.

An examination of the literature on the effectiveness of cochlear implants in prelingually deaf children suggests that "central" auditory, cognitive and linguistic factors may be responsible for some of the variability and individual differences observed in traditional outcome measures (see Pisoni, 2000; Pisoni et al., 2000). Although the NIH Consensus Statement on cochlear implants mentioned "central" auditory factors, the report was not very specific about precisely what these factors might be or what role higher-level cognitive processes might play in outcome measures.

In this chapter, we present a summary of recent findings that suggest that the observed individual differences in outcome following implantation are related to central cognitive factors associated with perception, attention, memory, learning and language processing. These new findings are encouraging and suggest additional directions for future research on the effects of early sensory experience and language development in children following cochlear implantation. New process measures of performance have revealed the contribution of working memory, rehearsal and coding processes to several outcome measures of speech and language development (e.g., measures of open-set word recognition and tasks requiring the use of phonological processing skills). Recent findings also suggest that the particular form of learning that occurs following cochlear implantation may be "domain-specific" and may be related to processing sound sequences and coding speech signals into phonological representations in working memory. These phonological representations form the basic building blocks of spoken language processing that are used in word recognition, comprehension and speech production.

Effectiveness of Cochlear Implants: Five Key Findings

Five key findings have been consistently reported in the literature on cochlear implants in deaf children. These findings suggest that the investigation of central auditory factors may provide new insights into the enormous variability in outcome and benefits observed following cochlear implantation. Knowledge and understanding of the cognitive factors responsible for these individual differences should be useful in helping patients obtain optimal benefits from their cochlear implants. In this section, we briefly review the five key findings that serve as the starting point for the research presented in this chapter. Then we summarize the major assumptions of the information processing approach to cognition that has guided our research program on developing new process measures of performance.

Individual Differences in Outcome and Benefit

Large individual differences in outcome and benefit following implantation are well documented in the clinical literature. However, all current outcome measures of performance are the final end product of a large number of complex sensory, perceptual, cognitive and linguistic processes that contribute to the observed variation among cochlear implant users. Until our recent studies on working memory in pediatric implant users, no research had focused on "process" or examined the underlying psychological and cognitive processes used to perceive and produce spoken language. Understanding these central cognitive processes already has provided new insights into the basis of individual differences and may help in developing new intervention techniques that can be used with patients who are deriving only minimal benefits from their implants. Four other findings have also been consistently reported in the literature on cochlear implants in children. These findings place several additional constraints on the problem of individual differences following cochlear implantation.

Age of Implantation and Length of Deafness

Age of implantation and length of deafness have both been found to affect a range of outcome measures. Children who receive an implant at an early age do consistently better on all of the clinical outcome measures than children who are implanted at older ages. Moreover, children who have been deprived of sound stimulation for shorter periods of time also do much better on a variety of outcome measures than children who have been deaf for longer periods of time. Both findings – age of implantation and length of deafness -- demonstrate the role of sensitive periods in development and the close links between neural development and behavior, especially the sensory and cognitive processes underlying hearing, speech and language development (Ball & Hulse, 1998; Konishi, 1985; Konishi & Nottebohm, 1969; Marler & Peters, 1988).

Effects of Early Experience

The nature of the early sensory and linguistic experience after implantation also has been found to affect performance on a wide range of outcome measures. Implanted children in "Oral-only" communication environments do much better on standardized tests of speech, language and vocabulary development than implanted children in "Total Communication" programs (Kirk, Pisoni, & Miyamoto, 2000). The differences in performance between these two groups of children as a function of "communication mode" are seen most clearly in receptive and expressive language tasks that make use of phonological processing skills such as open-set word recognition, language comprehension and measures of speech production, especially measures of a child's speech intelligibility and expressive language development (Cullington et al., 2000; Hodges et al., 1999; Kirk et al., in press; Svirsky, Sloan, Caldwell & Miyamoto, 2000).

Lack of Preimplant Predictors of Outcome

Until recently, researchers have been unable to identify any reliable behavioral preimplant predictors of outcome and success in children with cochlear implants (see Tait, Lutman & Robinson, 2000). The lack of reliable preimplant predictors in children is an important finding because it suggests the operation of complex interactions between the newly acquired sensory and perceptual capabilities of a child after a period of sensory deprivation, properties of the language-learning environment and the interactions with parents and caregivers that the child is exposed to early on after receiving a cochlear implant. More importantly, the absence of reliable preimplant predictors of outcome makes it difficult to identify in a timely manner, those children who may benefit from specific interventions to improve their speech and language processing skills.

"Emergence" of Abilities After Implantation

When all of the outcome and demographic measures are considered together, the current evidence suggests that the underlying sensory and perceptual abilities for speech and language "emerge" after implantation and that performance with a cochlear implant improves over time. Because the outcome and benefit of cochlear implantation cannot be predicted reliably from current pre-implant behavioral measures, improvement in performance observed after implantation is assumed to be due to learning and memory processes that are related in complex ways to maturational changes in neural and perceptual development and exposure to the target language in the child's immediate environment.

Taken together, the effects of demographic variables on outcome measures after cochlear implantation suggest several general conclusions about how cochlear implants facilitate the acquisition and development of spoken language. The findings also point to several underlying factors that may account for individual differences on various outcome measures. Although some proportion of the total variance in outcome performance is clearly due to peripheral factors related to audibility and the initial sensory encoding of the speech signal into "information-bearing" sensory channels in the auditory nerve, additional sources of variance may also come from more central "cognitive" factors. These additional sources of variance have to do with information processing operations and cognitive demands—that is, how the child uses the initial sensory input that is received from the cochlear implant and how the language-learning environment modulates, shapes and facilitates this learning process. Investigation of the encoding, rehearsal, storage and retrieval of information may provide some new insights into the underlying basis of the large individual differences in outcome measures of speech and language development.

To gain a better understanding of what deaf children are learning via their cochlear implants and how they use sound input, we have adopted a different theoretical perspective which looks more closely at the content and flow of information within the nervous system and how it changes over time following implantation. Our research on cochlear implants in children focuses on the underlying psychological and linguistic processes that mediate speech perception and production (see Pisoni, 2000).

Little, if any, of the previous research on cochlear implants (CIs) has explored what the children learn after they receive their implants, how they go about the process of acquiring language or how they develop receptive and expressive language skills. Until recently, there have been very few attempts to study the process of language development in deaf children with CIs and compare their linguistic knowledge and performance with that of normal-hearing children or hearing-impaired children who use hearing aids (Miyamoto et al., 1997; Robbins & Kirk, 1996).

These are important new research directions that go beyond the basic questions of clinical assessment, device efficacy, and measuring benefit with traditional outcome measures; they are fundamental problems in speech and hearing sciences that deal with the "effectiveness" of cochlear implants outside the restricted conditions of the hearing clinic or the research laboratory. The exclusive reliance on assessment-based clinical research and prediction of outcome measures following cochlear implantation has changed over the last few years. Several recent papers have already reported new findings on some of these issues (Kirk et al., 1997; Pisoni, Svirsky, Kirk, & Miyamoto, 1997; Robbins et al., 1998; Zwolan et al., 1997) and other studies are currently underway at a number of research centers around the world.

Information Processing Approach to Cognition

In order to pursue these new research questions and to move beyond the study of demographics and issues surrounding clinical assessment and prediction of outcome measures, it has become necessary to look to other allied disciplines for guidance. New experimental methods and behavioral techniques are available to study individual differences and the emergence of fundamental underlying cognitive and neural processes and how these change over time after implantation. Many useful experimental procedures already have been developed by cognitive and developmental psychologists to study perception, attention, learning and memory in children within the framework of human information processing (Haber, 1969; Lachman et al., 1979; Neisser, 1967). More importantly, this theoretical approach has also provided a variety of conceptual tools for thinking about the structures and processes involved in cognitive activity and the underlying psychological phenomena (Lindsay & Norman, 1977; Reitman, 1965).

Learning, Memory and Cognition Viewed as Information Processing

Information Processing Approach

The foundational assumption of our approach to understanding and explaining the variation and individual differences in speech and language outcome measures has been to view the human nervous system as an information processor. An information processor is a system that encodes, stores and manipulates various types of symbolic representations. Information can exist in several different forms at a number of levels of representation in the system, ranging from early registration and encoding of the sensory input to permanent storage of symbolic representations in long-term memory.

By viewing human cognition and traditional areas of basic research such as sensation, perception, attention, memory and learning as information processing within a larger integrated framework, cognitive scientists have obtained several important benefits. These include the development of new tools and experimental methodologies to study the processes that underlie these behaviors as well as the availability of a new theoretical conceptualization that can be used to explain and predict variability in complex higher-level behaviors such as speech and language in different clinical populations. The information processing approach to human cognition has also provided the theoretical motivation for reformulating some long-standing problems as well as identifying new research questions that can be studied within this framework. These research efforts have provided new insights into human performance and the neural and cognitive processing mechanisms that underlie these different behaviors.

Although there are many good clinical reasons to study prelingually deaf children who have received cochlear implants and to understand the basis for the variability in outcome measures of their speech and language skills, there are also several additional reasons to carry out research with this unique population that touch on basic theoretical issues related to neural development and behavior. For a variety of moral and ethical reasons, it is not possible to carry out sensory deprivation experiments with young children and it is not possible to delay or withhold treatment for an illness or disability that has been identified and diagnosed. Thus, for studies of this kind which are concerned with investigating the effects of early sensory experience on neural and behavioral development, it is necessary to rely on clinical populations who are receiving interventions of various kinds and hope that appropriate experimental designs can be developed which will yield new scientific knowledge.

Among the broader theoretical questions on neural and cognitive development that this research addresses are the following: What effect does the absence of sound and auditory stimulation during the first few years of life have on the development of the basic neural information processing mechanisms and skills used in speech and language processing? What effect does the introduction of sound and auditory stimulation by means of electrical hearing via a cochlear implant have on the development of speech and language processing after a period of auditory deprivation? What kind of a linguistic system (i.e., grammar) does a deaf child develop after receiving a cochlear implant? Is a deaf child's language delayed but otherwise typical of normal-hearing children or is it disordered or impaired in some fundamental way relative to normal-hearing, typically-developing, age-matched peers? These are a few of the theoretical questions this work focuses on.

Looking at the "Stars" – Analysis of the Exceptionally Good Implant Users

Several years ago, we carried out an analysis of a data set from a longitudinal project on cochlear implants in children at the Indiana University School of Medicine to gain some new insights into the basis of individual differences and variation in outcome in this clinical population (see Pisoni et al., 1997). We began first by looking at the "exceptionally" good users of cochlear implants—the so-called "Stars." These are the children who did extraordinarily well with their cochlear implants after only two years of implantation. The "Stars" appeared to acquire spoken language relatively quickly and easily and seemed to be on a developmental trajectory that paralleled normal-hearing children. At first glance, they look like normal hearing and normally developing children who simply have language delays (see Svirsky et al., 2000a, b).

To learn more about the "Stars," we analyzed outcome data from the children who scored exceptionally well on the PBK test two years after receiving their implant. The PBK test is an "open-set" test of spoken word recognition (see Meyer & Pisoni, 1999). Among clinicians, the PBK test is considered to be very difficult for prelingually deaf children compared to other, "closed-set" perceptual tests that are routinely used in standard assessment batteries (Kirk, Pisoni & Osberger, 1995; Zwolan, Zimmerman-Phillips, Asbaugh, Hieber, Kileny & Telian, 1997). Open-set tests of speech perception measure word recognition and lexical discrimination and require the child to search and retrieve the phonological representations of the test words from their lexical memory. These particular types of word recognition tests are extremely difficult for hearing-impaired children and adults with cochlear implants because the procedures and task demands require the listener to perceive and encode fine phonetic differences based entirely on information present in the speech signal without the aid of any external context or retrieval cues. Basically, the listener is required to discriminate, select and then identify a unique phonological pattern from a very large number of equivalence classes in lexical memory (see Luce & Pisoni, 1998).

The PBK score was used as the "criterial variable" to initially identify and select two groups of children for subsequent analysis using an extreme groups design. One group consisted of children who were exceptionally good cochlear implant users, the so-called "Stars." These were children who scored in the top 20% on the PBK test. A second set of children were selected as a comparison group. The children in this group scored in the bottom 20% on the PBK test and were unable to recognize any of the test words when they were presented in isolation using an open-set format.

After these children were selected and sorted into two groups based on PBK scores, we examined their performance on a variety of other outcome measures already obtained from them as part of our large-scale longitudinal project. These outcome measures included tests of speech perception, language comprehension, word recognition, receptive vocabulary knowledge, receptive and expressive language development and speech intelligibility scores. Descriptive analyses were carried out first to compare differences between the two groups on these measures. Then correlations were computed among the various outcome measures to look at relationships and commonalities between the measures.

The results of our descriptive analyses after one year of implant use revealed several interesting findings about the exceptionally good users of cochlear implants. First, we found that although the "Stars" showed better performance on some outcome measures such as speech perception, language comprehension, spoken word recognition and speech intelligibility than the control group, the two groups of children did not differ from each other on measures of receptive vocabulary knowledge, non-verbal intelligence, visual-motor integration or visual attention (see Pisoni et al., 1997). We also found that some outcome measures of performance continued to improve over the course of six years, whereas other outcome measures remained fairly stable after the first year.

These findings demonstrate that the "Stars" differ in selective ways from the comparison group of control subjects. Whatever differences are revealed by other descriptive measures, it is clear that the results are not due to some global difference in overall performance levels between the two groups. More importantly, we found that the "Stars" also displayed exceptionally good performance on another test of spoken word recognition, the LNT (Kirk et al., 1995), demonstrates that the superior lexical discrimination skills of these children are not due to the specific words on the PBK test or the particular methods used to administer the test. Instead, the differences appear to be related to a common set of information processing operations and procedures that are used by these children to carry out open-set word recognition tasks. Among the component elementary information-processing skills needed for this task are encoding, storage, rehearsal, imitation and speech production. These are the same basic skills that are used in all of the traditional clinical outcome measures used to assess performance in these children after implantation.

The results of our correlational analyses of the test scores for the "Stars" one year after implantation revealed a consistent pattern of strong and significant intercorrelations among several of the dependent variables, particularly measures of word recognition, language development and speech intelligibility, suggesting a common underlying source of variance that is shared by these measures (see Pisoni et al., 1997). The same patterns of intercorrelations were not observed for the comparison group. One common source of variance found in the correlational analyses of the "Stars" was related to the processing of spoken words and to the encoding, storage and rehearsal of the phonological representations of words.

Of particular interest was the unexpected finding of strong correlations of several outcome measures with speech intelligibility scores obtained for these children, suggesting transfer of knowledge between speech perception and production and the use of a common shared representational system for receptive and expressive language functions (see also Shadmehr & Holcomb, 1997). The results suggested that the exceptionally good performance of the "Stars" may be due to their superior spoken language processing abilities, specifically, their ability to perceive, encode and retrieve phonological representations of spoken words from lexical memory and use these linguistic representations in a variety of different language processing tasks, especially tasks that depend on the decomposition and re-assembly of the sound patterns of spoken words such as lexical retrieval, rehearsal and speech production. Our working hypothesis was that this particular source of variance might reflect "modality-specific" elementary information processing operations that are involved in the phonological coding of sensory inputs and the construction of phonological representations of spoken language.

While the results of our initial correlational analyses point to several new directions for future research on individual differences, the data available on these children were based on traditional outcome measures that were collected as part of the annual assessments in our longitudinal study. All of the scores on these tests are "endpoint measures" of performance, and as such they reflect the final product of perceptual and linguistic analysis. Process measures of performance, that is, measures of what a child does with the sensory information provided by his cochlear implant, were not part of the standard research protocol so it was impossible to investigate differences in speed, fluency or processing capacity. It is very likely that fundamental differences in neural and cognitive information processing may underlie the individual differences observed between the two groups of children in our initial study.

The analyses we carried out on the speech perception, word recognition, spoken language comprehension, vocabulary knowledge and language development scores revealed that a child who displayed exceptionally good performance on the PBK test also showed very good scores on a variety of other speech and language measures as well. We consider these new results to be theoretically important. The differences in outcome measures observed between the two groups suggested that it may be possible to determine more precisely how the "Stars" differ from the minimal benefit cochlear implant users. Knowledge of the factors that are responsible for individual differences in performance among deaf children who receive cochlear implants, particularly the variables that underlie the extraordinarily good performance of the "Stars," may be useful in helping the children who are not doing as well with their implants at an earlier point in development after implantation. Moreover, research on individual differences may have direct clinical relevance in terms of intervention in recommending specific changes to the child's language-learning environment and in modifying the nature of the interactions a child has with his parents, teachers and speech therapists, who provide the primary language model for the child. Research on individual differences may also help by providing clinicians and parents with a principled basis for generating realistic expectations about outcome measures, particularly measures of speech perception, comprehension, language development and speech intelligibility.

The results of our analyses of the "Stars" suggested several hypotheses about the source of the individual differences in performance. The primary locus of the differences in performance on many of the speech and language-based outcome measures may be due to central rather than peripheral factors. That is, the source of the individual differences may be related to how the initial sensory information is encoded, stored, retrieved and manipulated in various kinds of information processing tasks, such as speech feature discrimination, spoken word recognition, language comprehension and speech production.

One of the key components that links these processes together and serves as the "interface" or "gateway" between the initial sensory input and stored knowledge in long-term memory is working memory. Investigations of the properties of working memory may provide new insights into the nature and locus of the individual differences observed among users of cochlear implants (see Baddeley, Gathercole, & Papagno, 1998; Carpenter, Miyake & Just, 1994; Gupta & MacWhinney, 1997). Unfortunately, at the time our analyses of the "Stars" were carried out, we did not have any memory data available to test this hypothesis. Since that time, however, several new studies have been carried out with our collaborators at the Central Institute for the Deaf (CID) in St. Louis to study working memory in children with cochlear implants (Pisoni & Geers, 2000). The results of these experiments are reported in the sections below.

At a superficial level, it seems reasonable to conclude that the children who "hear" better through their cochlear implants simply learn language better and, subsequently, recognize words better. This generalization might be appropriate as an explanation of the perceptual data obtained from outcome measures of receptive function, but it is much more difficult to explain the observed differences in speech intelligibility and expressive language simply on the basis of better hearing and language-processing skills without a more detailed description of the underlying linguistic skills and abilities used in speech production.

To account for the differences in speech intelligibility and expressive language development, it is necessary to assume that a given child has acquired an underlying linguistic system that mediates between speech perception and speech production. Without assuming a common linguistic system—a grammar—we would have no reason to expect a child's receptive and expressive language abilities to be as closed coordinated as they typically are in normal-hearing, normal-developing children. It is well known that reciprocal links exist between speech perception, production and a whole range of language-related abilities; these interconnections reflect the child's linguistic knowledge of phonology, morphology and syntax. Speech perception, spoken word recognition and language comprehension are not isolated, autonomous perceptual abilities, independent of the child's developing linguistic system. Thus, explanations of the superior performance of the "Stars" framed in terms of hearing, audibility or sensory discrimination abilities cannot provide a satisfactory theoretical account of all of the results reported above or an adequate description of how early auditory experience affects speech perception and language development in these children. Some other non-sensory process must be responsible for the commonalities in speech perception and production observed across these diverse outcome measures.

In order to understand and explain the individual differences in these children, several additional performance measures are needed to assess how deaf children with cochlear implants process, code and represent the sensory, perceptual and linguistic information they receive through their implants, and how they use this information in a variety of behavioral tasks. The traditional endpoint outcome measures in our database were scores on behavioral tests used for assessment of specific speech and language skills thought to be important for measuring change and benefit after implantation. This battery of clinical tests was designed and constructed many years ago when theoretical issues about individual differences in performance were not high research priorities. As a result, no data were ever collected on cognitive processes such as memory, learning, attention, automaticity or modes of processing.

Some Measures of Information Processing Capacity

To obtain some initial measures of working memory capacity from a large number of deaf children following cochlear implantation, we were very fortunate to be able to collaborate with Ann Geers and her colleagues at CID, who already had an on-going large-scale research project underway. Their project was designed to obtain a wide range of outcome measures of speech, language and reading skills from 8 and 9 year old children who had all used their cochlear implants for at least three and one-half years. Thus, chronological age and length of implant use in the children were controlled in this study.

Using the test lists and procedures from the WISC III (Wechsler, 1991), forward and backward auditory digit spans were obtained from four groups of 8- and 9- year old deaf children with cochlear implants. A total of 176 children were tested in separate groups at CID during the summers of 1997, 1998, 1999 and 2000. Forward and backward digit spans were also collected from an additional group of

45 age-matched normal-hearing 8- and 9- year old children. These children were tested in Bloomington, Indiana, and served as a comparison group.

The WISC-III memory span task requires the child to repeat back a list of digits as spoken live-voice by an experimenter at a rate of approximately one digit per second (WISC-III Manual, Wechsler 1991). In the "digits-forward" section of the task, the child is required to simply repeat back the list as heard. In the "digits-backward" section of the task, the child is told to "say the list backward." In both parts of the WISC task, the lists begin with two items, and are increased in length upon successful repetition until a child gets two lists incorrect at a given length, at which point testing stops. Points are awarded for each list correctly repeated with no partial credit. The task was administered using live-voice presentation with the face of the clinician visible to the child.

A summary of the digit span results for all five groups of children is shown in Figure 1. Forward and backward digit spans are shown separately for each group. The children with cochlear implants are shown in the four panels on the left separately by year of testing; the normal-hearing children are shown on the right. Each child's digit span in points is calculated by summing the number of lists correctly recalled at each list length. The points score for forward digit span can vary between zero and 16; the points score for backward digit span can vary between zero and 14.

Inspection of the data shown in Figure 1 reveals an orderly and systematic pattern of the forward and backward digit spans for the deaf children with cochlear implants. All four groups are quite similar to each other; in each group, the forward digit span is longer than the backward digit span. The pattern is quite stable over the four years of testing despite the fact that these scores are based on separate independent groups of subjects. The difference in span length between forward and backward report was highly significant for the entire group of 176 deaf children and for each group taken separately ($p < .001$).

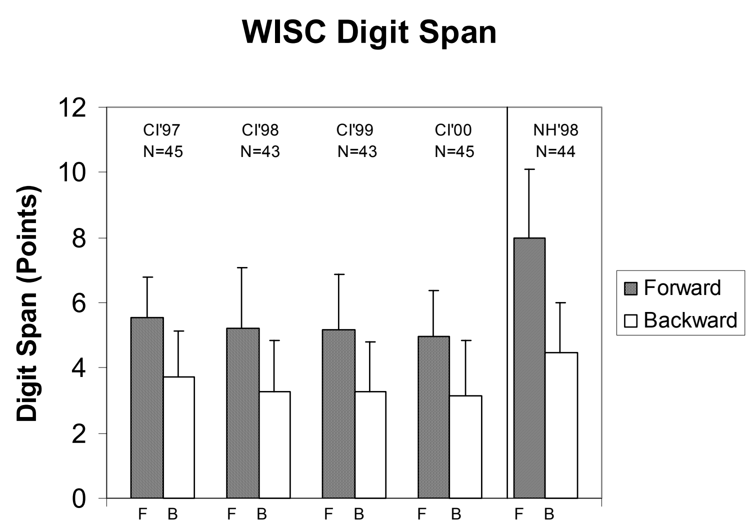


Figure 1. WISC digit spans scored by points for the four groups of 8- and 9-year old children with cochlear implants and for a comparison group of 8- and 9-year-old normal-hearing children. Forward digit spans are shown by the shaded bars, backwards digit spans by the open bars. Error bars indicate one standard deviation from the mean.

The forward and backward digit spans obtained from the 44 age-matched normal-hearing children are shown in the right-hand panel of the figure. Examination of these data shows that the digit spans for the normal-hearing children differ in several systematic ways from the digit spans obtained from the children with cochlear implants. First, both digit spans are longer than the spans obtained from the children with cochlear implants. Second, the forward digit span for the normal-hearing children is much longer than the forward digit spans obtained from the children with cochlear implants. This latter finding is particularly important because it suggests atypical development of the deaf children's short-term memory capacity and points to several possible differences in the underlying processing mechanisms that are used to encode and maintain sequences of spoken digits in immediate memory.

Numerous studies over the years have suggested that forward digit spans can be used to index and assess initial coding strategies related to phonological processing and verbal rehearsal mechanisms used to maintain information in short-term memory for brief periods of time before retrieval and output response. In contrast, differences in backward digit spans are thought to reflect the contribution of controlled attention and operation of "executive" processes used to recode, transform and manipulate verbal information for later processing operations (Rosen & Engle, 1997; Rudel & Denckla, 1974)

The digit spans for the normal-hearing children shown in Figure 1 are age-appropriate and fall within the published norms for the WISC III. However, the forward digit spans obtained from the children with cochlear implants are atypical and suggest possible differences in encoding and/or rehearsal processes used in immediate memory. In particular, the forward digit spans reflect possible differences in processing capacity of immediate memory between the two groups of children. These differences may cascade and affect other information processing tasks that make use of working memory and rehearsal processes. Because all of the behavioral tasks typically used to assess clinical speech and language outcomes following implantation make use of the component processes of working memory and rehearsal, it seems reasonable to assume these tasks will also reflect variation due to individual differences in working memory and processing capacity.

Correlations with Digit Spans

In order to learn more about the observed differences in auditory digit span and the limitations in processing capacity in the children with cochlear implants, we examined the correlations between forward and backward digit spans and several speech and language outcome measures that were obtained from the same children at CID. Of the various demographic measures available, the only one that correlated strongly and significantly with digit span was a measure called "Communication Mode." This measure is used to quantify the nature of the child's early sensory and linguistic experience after receiving a cochlear implant in terms of the degree of emphasis on oral language skills by parents, teachers and therapists in the home and educational environments.

Each child's degree of exposure to Oral-only communication methods was quantified by determining the type of communication environment experienced by the child in the year just prior to implantation, each year over the first three years of CI use, and then in the year just prior to the current testing. A score was assigned to each year, ranging from a "1," corresponding to the use of "total communication" with a sign emphasis (that is, indicating extensive use of manual signs in addition to spoken language), to "6," indicating an auditory-verbal environment with a strong emphasis on auditory communication without the aid of lipreading (see Geers et al., 1999 for details). Communication methods intermediate between these two extremes were assigned intermediate scores ranging from 2 to 5. These scores were then averaged over the five points in time. The mean communication mode score for the group over the five intervals was approximately 3.9 on this 6-point scale. However, a wide range of communication mode backgrounds was present within the sample of 176 children (range of average communication mode scores = 1.0 to 6.0).

We found that forward digit span was positively correlated with Communication Mode ($r = +.34$, $p < .001$); children in language learning environments that primarily emphasized oral skills displayed longer forward digit spans than children who were in total communication (TC) environments. However, the correlation between digit span and communication mode was highly selective in nature, because it was restricted only to the forward digit span scores; the backward digit spans were not correlated with communication mode or any other demographic variable.

In order to examine the effects of early experience in greater detail, a median split was carried out on the communication mode scores to create two subgroups, Oral children and TC children. Figure 2 shows the digit spans plotted separately for the Oral and TC children for each of the four years of testing at CID. Examination of the forward and backward digit spans for these two groups of children indicates that the Oral group consistently displayed longer forward digit spans than the TC group. While the differences in forward digit span between Oral and TC groups were highly significant, the differences in backward digit span were not. This pattern suggests that the effects of early sensory and linguistic experience on immediate memory is selective in nature and appears to be restricted to coding and rehearsal processes that affect only the forward digit span task.

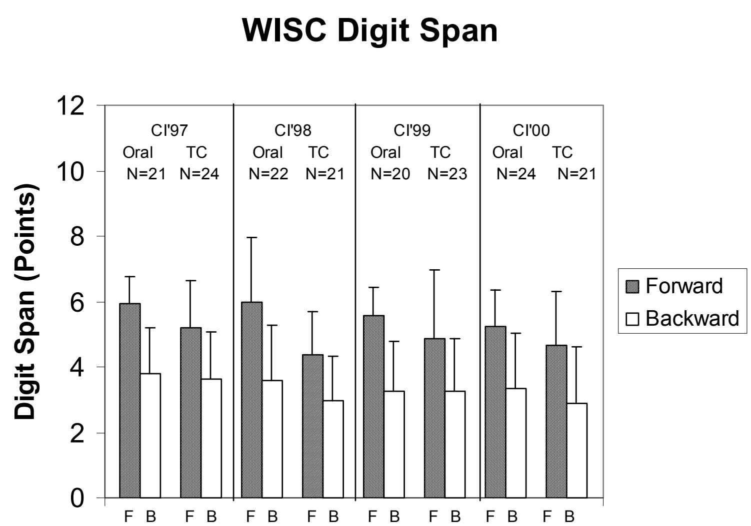


Figure 2. WISC digit spans scored by points for the four groups of 8- and 9-year old children with cochlear implants, separated by communication mode. For each year, scores for the oral group are shown to the left of those for the total communication group. Forward digit spans are shown by the shaded bars, backwards digit spans by the open bars. Error bars indicate one standard deviation from the mean.

The difference in forward digit span between Oral and TC children present for each of the four groups suggests that forward digit spans are sensitive to the nature of the early sensory and linguistic experience that the child receives immediately after cochlear implantation. The differences observed in the forward digit spans could be due to several factors, such as more efficient encoding of the initial stimulus patterns into more stable phonological representations in working memory, speed and efficiency of the rehearsal processes that are used to store and maintain information in working memory or possibly even speed of scanning and retrieval of information in working memory after recognition has taken place. All three factors could influence measures of processing capacity and any one of these could affect the number of digits correctly recalled from immediate memory in this task.

Regardless of which factor or factors are responsible for the differences observed above, these results demonstrate that forward digit span is sensitive to the effects of early sensory and linguistic experience and suggest that several specific mechanisms in the information processing system may be affected by the nature of the early experience the child receives after implantation. Although these results clearly demonstrate that early experience in an environment that emphasizes oral language skills is associated with longer digit spans and increased information processing capacities of working memory, without additional converging measures of performance, it is difficult to specify precisely what elementary information processing mechanisms are actually affected by early experience and which ones are responsible for the increases in forward digit spans observed in these particular children.

Digit Spans and Word Recognition

As mentioned above in the summary of our findings on the "Stars," the large individual differences we observed in a range of outcome measures of speech and language development appear to be related in some way to spoken word recognition skills and to tasks that make use of phonological representations of spoken words. At the time our research on the "Stars" was carried out, we did not have digit span data or any other process measures of performance from this group of deaf children. We assumed that the word recognition skills of the "Stars" would draw on the same basic component processes that are used in open-set tests of word recognition like the PBK test, which was used as the criterion variable to identify the exceptionally good implant users and sort our subject population into two extreme groups. While a number of demographic factors such as duration of deafness, length of device use and age at implantation have been shown to be related to variation in these outcome measures, these variables are only able to account for a small portion of the observed variance in performance. We reasoned that some other factor must be responsible for the wide range of performance observed in these deaf children.

Our current hypothesis is that a portion of the remaining unexplained variance can be accounted for in terms of individual differences in the elementary information processing operations that are related to the speed and efficiency with which phonological representations of spoken words are maintained and retrieved from memory after recognition and identification has taken place. Numerous studies of normal-hearing children over the past few years have demonstrated close links between verbal short-term memory and learning to recognize and understand new words (Baddeley, Gathercole, & Papagno, 1998; Gupta & MacWhinney, 1997). Other research has found that vocabulary development and several important milestones in speech and language acquisition are also associated with differences in measures of verbal short-term working memory (e.g., specifically, measures of digit span), which can be used as estimates of processing capacity.

To determine if measures of digit span and capacity of immediate memory are related to spoken word recognition in deaf children following cochlear implantation, we examined the correlations between the WISC forward and backward digit span scores and three different measures of word recognition that were also obtained from the children tested at CID in 1997 and 1998. A summary of the correlations between digit span and word recognition scores based on these 88 children is shown in Table I for the WIPI, LNT and BKB word recognition tests.

Table I

	Simple Bivariate Correlations		Partial Correlations[a]	
	WISC Forward Digit Span	WISC Backward Digit Span	WISC Forward Digit Span	WISC Backward Digit Span
Closed Set Word Recognition (WIPI)	.50***	.33**	.31**	.13
Open Set Word Recognition (LNT-E)	.54***	.31**	.33**	.10
Open Set Word Recognition in Sentences (BKB)	.56**	.37***	.35**	.12

*** $p < .001$, ** $p < .01$

[a]Statistically Controlling for Communication Mode Score, Age of Onset of Deafness, Duration of Deafness, Duration CI Use, Number of Active Electrodes, VIDSPAC Total Segments Correct (Speech Feature Perception Measure), Age

The WIPI (Word Intelligibility by Picture Identification Test) is a closed-set test of word recognition in which the child selects a word from among six alternative pictures (Ross & Lerman 1979). The Lexical Neighborhood Test (LNT) is an open-set test of word recognition and lexical discrimination that requires the child to imitate and reproduce an isolated word (Kirk et al., 1995). This test is similar to the well-known PBK test, although the vocabulary on the LNT was designed to control for familiarity while lexical competition among the items was manipulated systematically to measure discrimination among phonetically similar words in the child's lexicon. Finally, the BKB test is an open-set word recognition test in which key words are presented in sentence contexts (Bench, Kowal & Bamford, 1979).

Table I displays two sets of correlations. The left-hand portion of the table shows the simple bivariate correlations of the forward and backward digit spans with the three measures of word recognition. Examination of the correlations for both the forward and backward spans reveals that children who had longer WISC digit spans also displayed higher word recognition scores on all three tests. These correlations are all positive and reached statistical significance although the correlations of forward digit span with the word recognition scores are somewhat larger than the correlations found for the backward span.

The right-hand portion of the table shows a summary of the partial correlations among these same measures after statistically controlling for differences due to chronological age, communication mode, duration of deafness, duration of device use, age of onset of deafness, number of active electrodes and speech feature discrimination. When these seven other "contributing variables" were statistically removed from the correlational analyses, the partial correlations between digit span and word recognition scores became smaller in magnitude. However, the correlations of the forward digit span with the three word recognition scores are still positive and statistically significant while the correlations of the backward digit spans are weaker and no longer significant. These results demonstrate that children who have longer forward WISC digit spans show higher word recognition scores and that this relationship is observed for all three word recognition tests even after the other sources of variance are removed.

Forward digit span accounts for approximately 11% of the currently unexplained variance in the word recognition scores while the backward digit span accounts for only 1.4% of the variance in these scores. The present results suggest the presence of a common source of variance that is shared between forward digit span and measures of spoken word recognition that is independent of other obvious mediating factors that have been found to contribute to the variation in these outcome measures.

Digit Spans and Speaking Rate

While the correlations of the digit span scores with communication mode and spoken word recognition scores suggest fundamental differences in encoding and rehearsal speed which are influenced by the nature of the early experience a child receives, these measures of immediate memory span and estimates of information processing capacity are not sufficient on their own to identify the underlying processing mechanism (or mechanisms) that are responsible for the individual differences. Additional converging measures are needed to pinpoint the locus of these processing differences. Fortunately, an additional set of behavioral measures was obtained from these children for a different purpose and made available to us for several new analyses. These data consisted of a set of acoustic measurements of speech samples from each child. These speech samples provided a unique opportunity for us to use converging measures to understand and explain the digit span results.

As part of the on-going research project at CID, several speech production samples were obtained from each child in order to assess speech intelligibility and measure changes in articulation and phonological development following implantation (see Tobey et al., 2000). The speech samples consisted of three sets of meaningful English sentences that were elicited using the stimulus materials and experimental procedures developed by McGarr (1983) to assess the speech intelligibility and articulation of deaf children. All of the utterances produced by the children were originally recorded and stored digitally for playback to groups of naïve adult listeners who were asked to transcribe what they thought the children had said. In addition to the speech intelligibility scores that were obtained for each child using these playback procedures, we analyzed the duration measurements of the individual sentences in each set and used these measures as estimates of each child's articulation rate.

We knew from a large body of research in the memory literature that a child's articulation rate is closely related to speed of subvocal rehearsal (Cowan et al., 1998). Numerous studies in the literature over the past 25 years have demonstrated strong relations between speaking rate and memory span for digits and words. The results of these studies have been replicated with several different populations and suggest that measures of an individual's speaking rate reflect articulation speed, which in turn, can be thought of as an index of rate of covert rehearsal for verbal (phonological) materials in working memory (Baddeley, Thompson & Buchanan, 1975). Individuals who speak more quickly have been found to have longer memory spans than individuals who speak more slowly. Measures of speaking rate are assumed to reflect articulation speed, which in turn, has been taken as an index of verbal rehearsal speed in working memory. Thus, individuals who have faster rehearsal speeds tend to show longer memory spans for sequences of digits and words.

Several different explanations of these findings have been proposed. One account assumes that more forgetting occurs from immediate memory at slower speaking rates because fewer words can be articulated and perceived within the same period of time. Another proposal assumes that the mechanism that controls speaking rate is the same one that regulates the speed of verbal rehearsal processes in immediate memory. Thus, more words can be maintained at faster rehearsal speeds. Regardless of which view is correct, the relation observed between measures of speaking rate and immediate memory span is a reliable and robust finding in the literature on working memory that has been observed in several different populations of subjects.

The forward digit span scores for the 88 children tested in 1998 and 1999 are shown in Figure 3 along with estimates of their speaking rates obtained from measurements of the seven syllable McGarr sentences. The digit spans are plotted on the ordinate; the average sentence durations are shown on the abscissa. The top panel shows mean sentence durations; the bottom panel shows the logarithmic transformations of the sentence durations. The pattern of results in both figures is clear; children who produce sentences with longer durations speak more slowly and, in turn, have shorter forward digit spans. The correlations between forward digit span and both measures of sentence duration were strongly negative and highly significant (r = -.63 and r = -.70; p <.001, respectively). The simple bivariate correlations between forward digit span and both the raw and transformed measures of sentence duration were also strongly negative and highly significant (r = -.55 and r = -.59; p < .001, respectively). For backwards digit span, the observed correlations were somewhat smaller, but still statistically significant (r = -.42 and r = -.42; p < .001).

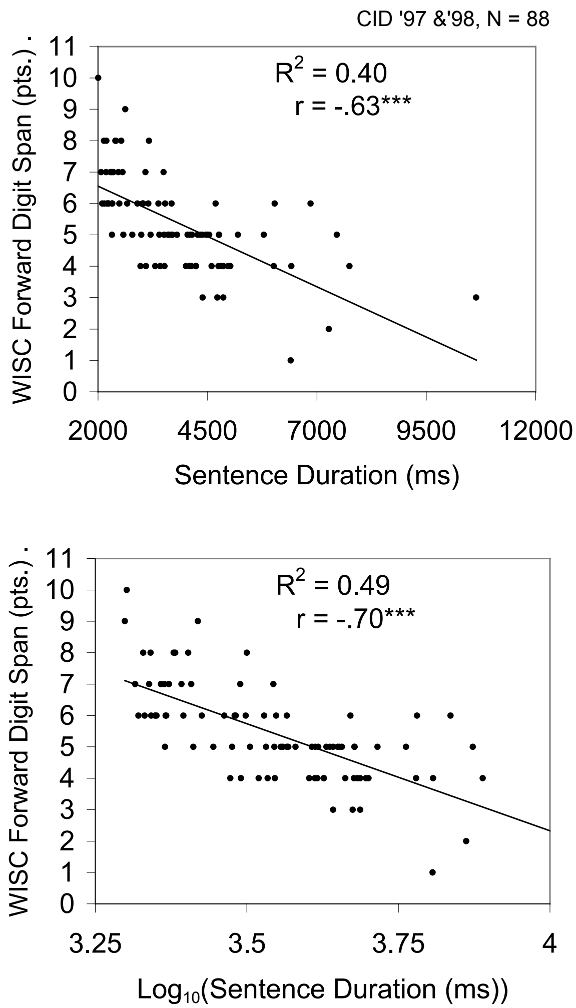


Figure 3. Scatterplots illustrating the relationship between average sentence duration for the seven-syllable McGarr Sentences (abscissa) and WISC forward digit span scored by points (ordinate). Each data-point represents an individual child. Non-transformed duration scores are shown in the top panel, log-transformed duration scores in the bottom panel. R-squared values indicate percent of variance accounted for by the linear relation.

These findings demonstrate that verbal digit span and articulation rate are correlated in this clinical population, as they are in normal-hearing school-age children and adults. That is, children who speak more quickly tend to have larger working memory capacities reflected by their longer digit spans. This result suggests the existence of a common information processing mechanism that is responsible for the individual differences observed within both tasks, namely, covert verbal rehearsal speed.

Speaking Rate and Word Recognition

To determine if rehearsal speed is also related to individual differences in word recognition performance, we examined the correlations between sentence durations and the three different measures of spoken word recognition described earlier. Table II shows the correlations between articulation rate and word recognition scores on the WIPI, LNT and BKB. All of these correlations are negative and strong, suggesting once again that a common processing mechanism, which we assume is related to verbal rehearsal speed, may be the primary factor that underlies the variation and individual differences observed in all three word recognition tasks.

Table II

	Simple Bivariate Correlations	
	Sentence Duration	Log (Sentence Duration)
Closed Set Word Recognition (WIPI)	-.65***	-.69***
Open Set Word Recognition (LNT-E)	-.59***	-.66***
Open Set Word Rec. in Sentences (BKB)	-.71***	-.78***

*** $p < .001$, ** $p < .01$

Our analysis of the digit span scores from these deaf children has uncovered two important correlations linking forward digit span to word recognition performance, on the one hand, and forward digit span to speaking rate (as indexed by measures of sentence duration), on the other hand. These correlations with forward digit span suggest the possibility of a common underlying processing factor that is shared by each of these dependent measures. This factor may reflect the speed of verbal rehearsal processes in working memory. If this hypothesis is correct, then word recognition and speaking rate should also be correlated with each other because they make use of the same processing mechanism. This is exactly what we found.

Table III shows a summary of the partial correlations computed between the two measures of speaking rate based on the McGarr sentence durations and the three measures of spoken word recognition performance described earlier. As in the earlier analyses, differences due to demographic factors and the contribution of other variables were statistically controlled for by computing partial correlations. In all cases, the correlations between speaking rate and word recognition were negative and highly significant. Thus, slower speaking rates as measured by longer sentence durations are associated with poorer word recognition scores on all three word recognition tests. Sentence duration accounted for approximately 25% of the currently unexplained residual variance in the word recognition scores after the other mediating variables were removed. These findings linking speaking rate and word recognition suggest that all three measures (digit span, speaking rate and word recognition performance) are related because they share a common underlying source of variance.

Table III

	Partial Correlations[a]	
	Sentence Duration	Log (Sentence Duration)
Closed Set Word Recognition (WIPI)	-.50***	-.55***
Open Set Word Recognition (LNT-E)	-.38***	-.47***
Open Set Word Rec. in Sentences (BKB)	-.52***	-.64***

*** $p < .001$, ** $p < .01$

[a]Statistically Controlling for Communication Mode Score, Age of Onset of Deafness, Duration of Deafness, Duration CI Use, Number of Active Electrodes, VIDSPAC Total Segments Correct (Speech Feature Perception Measure), Age

In order to determine if digit span and sentence duration share a common process and the same underlying source of variance which relates them both to word recognition performance, we re-analyzed the intercorrelations between each pair of variables with the same set of the demographic and mediating variables systematically partialled out. When sentence duration was partialled out of the analysis, the correlations between digit span and each of the three measures of word recognition approached zero. However, the negative correlations between sentence duration and word recognition were still present even after digit span was partialled out of the analysis suggesting that processing speed is the common factor that is shared between these two measures.

The pattern of results that emerges from these analyses suggests that the underlying process that is shared in common with sentence duration is related in some way to the rate of information processing, specifically, to the speed of verbal rehearsal in working memory. This processing component of verbal rehearsal could reflect either the actual articulatory speed used to recycle and maintain phonological patterns in working memory or the time to retrieve and scan items already in working memory (see Cowan et al. 1998). In either case, the common factor that links word recognition and speaking rate appears to be related to the speed of information processing operations used to maintain phonological information in working memory. Thus, variation in performance in these two tasks can be traced to a common elementary process that is shared by both measures of performance.

These new findings demonstrating a relation between speaking rate and digit span permit us to identify a specific information processing mechanism, the verbal rehearsal process, which appears to be responsible for the limitations on processing capacity. Processing limitations are present in a range of behavioral tasks that make use of verbal rehearsal and phonological processing skills to encode, store, maintain and retrieve spoken words from working memory. We suggest that these fundamental information-processing operations are common components of all current outcome measures that are routinely used to assess both receptive and expressive language functions. The present findings suggest that the variability in performance on the traditional clinical outcome measures used to assess speech and language-processing skills in deaf children after cochlear implantation may reflect fundamental differences in the speed of information processing operations such as verbal rehearsal and the rate of encoding phonetic and lexical information in working memory.

Simon Reproductive Memory Spans

The traditional methods for measuring working memory using digit spans all require a subject to verbally imitate and repeat back a sequence of test items using an overt articulatory response. Because most deaf children with cochlear implants also have delays in speech development and display "atypical" articulation and speech motor control, it is possible that any differences observed in working memory using digit spans could be due to the nature of the response requirements during retrieval and output in addition to possible differences in encoding, storage and rehearsal processes.

To eliminate the use of an overt articulatory-verbal response, we developed a new experimental methodology to measure memory spans based on Milton Bradley's Simon, a popular memory game. Figure 4 shows a display of the apparatus. In this procedure, a subject simply "reproduces" a stimulus pattern by manually depressing a sequence of colored response panels on a four-alternative response box. The Simon memory game procedure also permitted us to manipulate the stimulus presentation conditions in several ways while holding the response format constant. This particular property of the experimental procedure is quite useful in providing a way to measure how various perceptual dimensions of the visual and auditory modalities are analyzed and processed, alone and in combination. The Simon memory game apparatus and methodology also provided us with an opportunity to study learning, specifically, sequence learning and the relations between memory capacity and learning using the same experimental procedures and response demands.

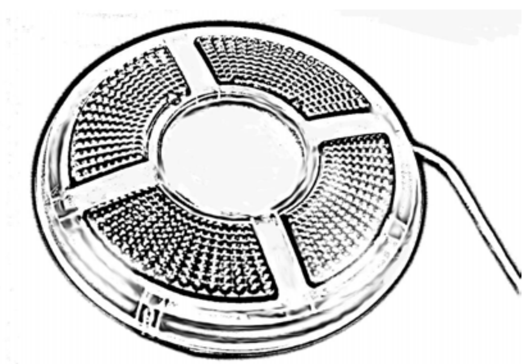
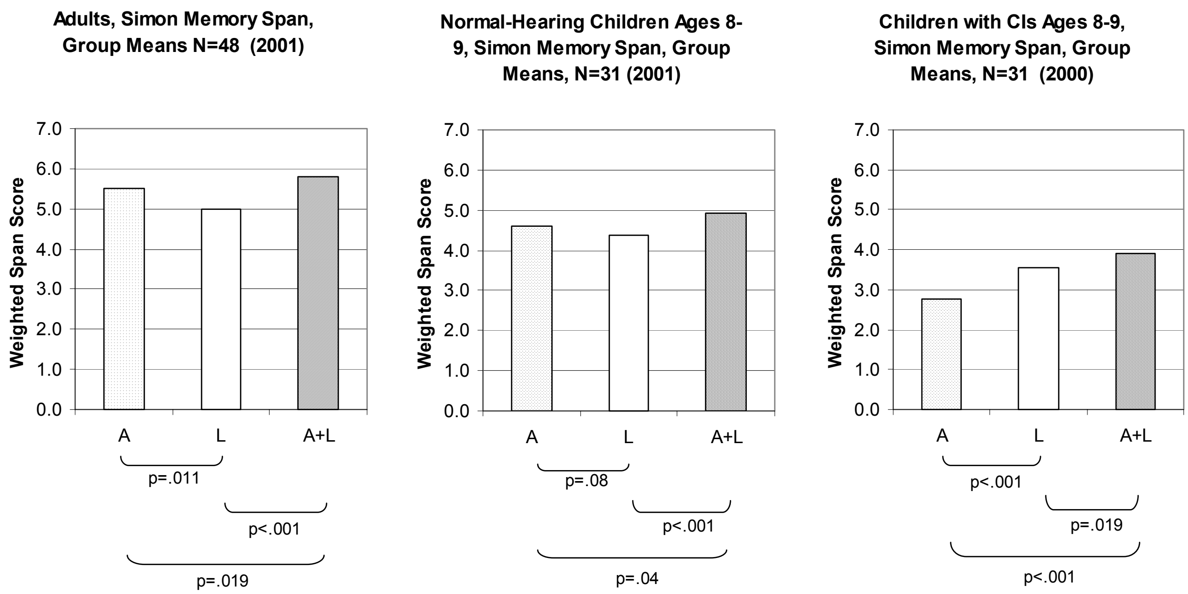


Figure 4. The Memory Game response box based on the popular Milton Bradley game "Simon."

The lights on the Simon are arranged in temporal patterns that systematically increase in length as the subject progresses through successive finals in the experiment. An adaptive testing algorithm is used to control presentation. If the child reproduces a pattern correctly twice in a row, the pattern increases in length on the next trial until the child is no longer able to reproduce a sequence correctly. Before the memory game was administered, each child was asked to identify the recorded tokens of the color-names by pointing to the four large colored buttons on the response box.

Sequences used for the memory game task were generated pseudo-randomly by a computer program, with the stipulation that no single item would be repeated consecutively in a given list. Each subject started with a list length of one item. If two lists in a row at a given length were correctly reproduced, the next list presented was increased by one item in length. If on any trial the list was incorrectly reproduced, the next trial used a list one item shorter in length. This "adaptive tracking procedure" is similar to methods typically used in psychophysical testing (Levitt, 1970). We computed a "weighted" span score for each child by finding the proportion of lists correct at each list length and summing these proportions across all list lengths.

A summary of the results from the Simon reproduction memory task for the three groups of subjects is shown in Figure 5. The normal-hearing adults are shown in the left panel, the normal-hearing aged-matched children are shown in the middle panel and the children with cochlear implants are shown in the right panel. Within each panel, the scores for auditory-only presentation (A) are shown on the left, scores for lights-only presentation (L) are shown in the middle and scores for the combined auditory and lights presentation condition (A+L) are shown on the right.



p-values from multiple paired samples t-tests, p-values uncorrected

Figure 5. Mean working memory spans in each of the three conditions tested using the "Simon" response box. Scores for a group of normal-hearing adults are shown on the left, scores for normal-hearing 8- and 9-year-old children are shown in the center, and scores for a group of 8- and 9-year-old cochlear implant users are shown on the right. Speckled bars indicate mean spans in the auditory-only (A) condition, open bars indicate mean spans in the lights-only (L) condition, and shaded bars indicate mean spans in the auditory-plus-lights (A+L) condition.

Examination of weighted Simon memory span scores for the normal-hearing adults reveals several findings that can serve a useful benchmark for evaluating differences in performance of the other two groups of children. First, we found a "modality effect" in presentation format. Auditory presentation of sequences of color names produced longer memory spans than visual presentation of sequences of colored lights ($p < .01$). Second, we found a "redundancy gain." When information from separate auditory and visual modalities was combined together and presented simultaneously, the memory span scores increased compared to presentation using only one sensory modality ($p < .02$ for auditory-only and $p < .001$ for visual-only, respectively).

The modality effect and the redundancy gains demonstrate that the Simon memory game procedure is a valid methodology for measuring immediate memory span in normal-hearing adults because is able to assess subtle differences in the sensory modality used for presentation of the stimulus patterns. As in other studies of verbal short-term memory, longer Simon memory spans were found for auditory stimuli compared to visual stimuli, suggesting the use of phonological coding and verbal rehearsal strategies (Penny, 1989; Watkins, Watkins & Crowder, 1974). In addition, the Simon memory spans were sensitive to cross-modality redundancies between stimulus dimensions when the same information about a stimulus pattern was presented simultaneously to more than one sensory modality. This latter finding demonstrates that adults are not only able to combine and integrate redundant sources of information across different sensory modalities but they are also able to increase their working memory capacity when stimulus redundancies are present in both auditory and visual modalities simultaneously.

The middle panel of Figure 5 shows the results of the three presentation conditions for the group of normal-hearing 8-and 9-year old children who were age-matched to the group of deaf children who use cochlear implants. Overall, the pattern of the Simon weighted span scores is quite similar to the findings obtained with the normal-hearing adults (shown in the left-hand panel), although there are several differences worth pointing out. First, the absolute memory span scores for all three presentation conditions are lower than the scores obtained from the adults. Second, while the modality effect found with the adults is also present in these data, it is smaller in magnitude and marginally significant, suggesting possible developmental differences in the rate and efficiency of verbal rehearsal between adults and children in processing auditory and visual sequential patterns like those used in this task. The cross-modal "redundancy gain" observed with the adults was also found with the normal-hearing children, although it is also smaller in magnitude ($p < .04$ for auditory- only; $p < .001$ for visual-only, respectively). Again, these differences may simply be due to age, maturation and development.

The Simon memory span scores for the deaf children with cochlear implants are shown in the right-hand panel of Figure 5 for the same three presentation conditions. Examination of the pattern of these memory span scores reveals several important differences from the span scores obtained for the normal-hearing children. First, the memory spans for all three presentation conditions were consistently lower overall than the spans from the corresponding conditions obtained for the normal-hearing children. Second, the modality effect observed in both the normal-hearing adults and normal-hearing children is reversed for the deaf children with cochlear implants; the memory spans for visual-only presentation were longer than auditory-only presentation, and this difference was highly significant ($p < .001$). Third, although the cross-modal "redundancy gain" found for both the adults and normal-hearing children was also observed for the deaf children and was statistically significant for both conditions ($p < .001$ for auditory-only and $p < .02$ for visual-only), the size of the gain was much smaller. Moreover, the differences in the magnitude of the gain relative to performance in the auditory-only and visual-only conditions were also different because of the reversal of the modality effect in the deaf children.

The results shown in Figure 5 for the visual-only presentation conditions are of special theoretical interest because the deaf children with cochlear implants displayed shorter memory spans than the normal-hearing children. This was an unexpected finding that adds support to the hypothesis that recoding and verbal rehearsal processes in working memory may play an important role in perception, learning and memory in these children. Capacity limitations of working memory are closely tied to speed of processing information even for visual patterns which are rapidly recoded and represented in memory in a phonological or articulatory code for certain kinds of sequential processing tasks. Verbal coding strategies may be mandatory in memory tasks that require immediate serial recall of temporal patterns that preserve item and order information (Gupta & MacWhinney, 1997). Thus, although the visual patterns were presented using only sequences of lights, both groups of children may have attempted to recode the sequential patterns using verbal coding strategies to create stable phonological representations in working memory for maintenance and rehearsal prior to response output.

Although normal-hearing adults and normal-hearing children showed a similar pattern of memory span scores across the three presentation conditions, the deaf children may have used a different encoding strategy and different rehearsal processes for maintaining temporal sequences in working memory. Auditory deprivation and the resulting absence of sound stimulation due to deafness during early stages of development may affect not only early sensory processing and perception but also subsequent encoding and rehearsal processes in working memory. These deaf children showed a reduced capacity to maintain temporal information in working memory even when information that was initially presented through the visual sensory modality. These findings on working memory spans for auditory and visual patterns obtained with the Simon memory game, which did not require overt verbal articulatory-motor responses, are consistent with the earlier memory span results obtained using the WISC digit spans, which showed systematic differences between the deaf children with implants and normal-hearing children.

To our knowledge, these are the first memory span data collected from deaf children with cochlear implants demonstrating specific effects on working memory capacity and rehearsal processes without relying on an articulatory-based verbal response for output. Under all three presentation conditions, the children used the same manual responses to reproduce the stimulus sequences. The deaf children also showed much smaller redundancy gains in the multi-modal presentation conditions, which suggests that in addition to differences in working memory capacity and rate of verbal rehearsal, their information processing skills and abilities to perceive and encode multi-dimensional stimuli are atypical and somewhat compromised relative to age-matched normal-hearing children. The smaller redundancy gains observed in these deaf children may also be due to the reversal of the typical modality effect observed in studies of working memory that reflect verbal coding of the stimulus materials. The modality effect in short-term memory studies is generally thought to reflect phonological coding and verbal rehearsal strategies that actively maintain temporal order information of sequences of stimuli in immediate memory for short periods of time (see Watkins et al., 1974).

Simon Learning Spans

The first version of our Simon memory game used novel sequences of color names or colored lights. All of the sequences were generated randomly in order to prevent any learning from occurring, other than the routine adaptation that normally is observed in learning how to do a new task in a laboratory setting. Our primary goal was to obtain estimates of working memory capacity for temporal patterns that were not influenced by sequence repetition effects or idiosyncratic coding strategies that might increase memory capacity from trial to trial. Each test sequence was created on the fly by a random numbers generator so that the internal structure of a sequence of colors was always different and varied from trial to trial during the course of the experiment. If a subject correctly reproduced a pattern at a given length twice in a row, the adaptive testing algorithm in the experimental control program automatically increased the length of the sequence by one item on the next trial and then generated an entirely new temporal sequence of colors that was different from the sequence presented on the previous trial. This procedure was used throughout the entire experiment to obtain estimates of immediate memory capacity. Thus, there was no basis for learning to take place and the measures of Simon memory span can be used as estimates of capacity of immediate memory.

We have also used the same basic Simon memory game methodology to study sequence learning and to investigate the effects of long-term memory on coding and rehearsal strategies in working memory. To accomplish this goal and to be able to directly compare the gains in learning and the increases in working memory capacity to our earlier Simon memory span measures, we examined the effects of sequence repetition on immediate memory span by simply repeating the same pattern again if the subject correctly reproduced the sequence on a given trial. Thus, the same stimulus pattern was repeated over and over again on each trial for an individual subject and gradually increased in length by one item after each correct response until the subject was unable to reproduce the pattern correctly anymore. This provided an opportunity to study learning based on pattern repetition and to investigate how repetition affects the capacity of immediate memory.

Figure 6 displays a summary of the results obtained in the Simon learning conditions for the same three presentation formats used in the earlier conditions, that is, auditory-only (A), lights-only (L) and auditory+lights (A+L). The weighted memory span scores for the sequence learning conditions are shown on the right-hand side of each panel in this figure; the corresponding set of memory span scores obtained earlier under random presentation format for the same three presentation conditions are reproduced on the left-hand side of each panel. The data for the normal-hearing adults are shown in the left panel; the data for the normal-hearing 8-and 9-year old children are shown in the middle panel and the data for the deaf children with cochlear implants are shown in the right panel.

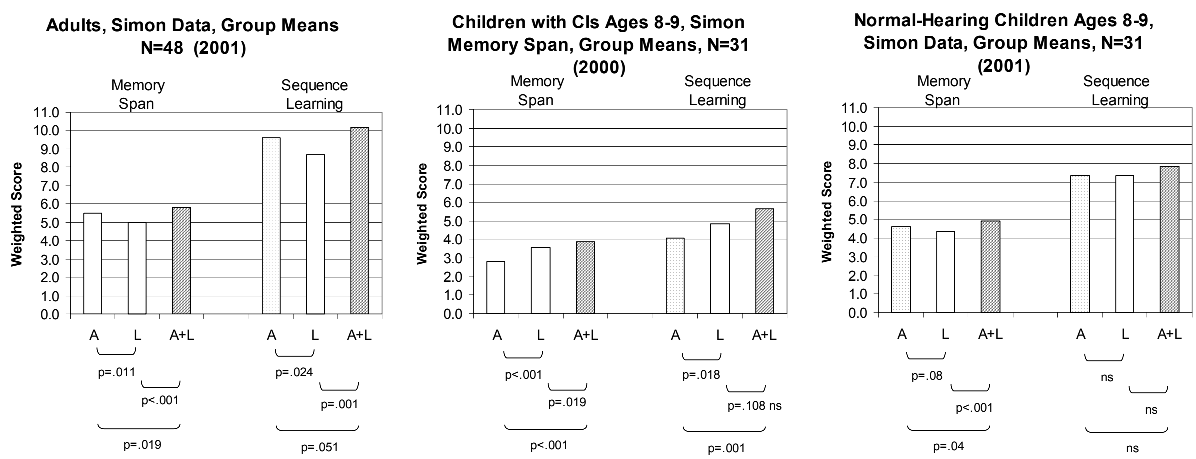


Figure 6. Mean working memory spans and mean sequence learning scores in each of the three conditions tested using the “Simon” response box. Scores for a group of normal-hearing adults are shown on the left, scores for normal-hearing 8- and 9-year-old children are shown in the center, and scores for a group of 8- and 9-year-old cochlear implant users are shown on the right. Speckled bars indicate mean scores in the auditory-only (A) condition, open bars indicate mean scores in the lights-only (L) condition, and shaded bars indicate mean scores in the auditory-plus-lights (A+L) condition. For each task, p-values for paired samples t-tests (uncorrected) between conditions are provided.

Examination of the two sets of memory span scores shown within each panel reveals several consistent findings. First, just repeating the same stimulus sequence again produced robust learning effects for all three groups of subjects. This repetition effect can be seen clearly by comparing the three scores on the right-hand side of each panel to the three scores on the left-hand side. In every case, the learning span scores are higher than the memory span scores; repetition of a pattern increased immediate memory span capacity, although the magnitude of the learning effects differed across the three groups of subjects. The memory spans observed for the adults in the learning condition are about twice the size of memory spans observed when the sequences were generated randomly from trial to trial. Although a repetition effect was also obtained with the deaf children who use implants, the size of their repetition effect was about half the size of the repetition effect found for the normal-hearing children shown in the middle panel.

Second, the rank ordering of the three presentation conditions in the sequence learning conditions was similar to the rank ordering observed in the memory span conditions for all three groups of subjects. The repetition effect was largest for the A+L conditions for all three groups. For both the normal-hearing adults and children, we also observed the same modality effect in learning that was found for memory span; auditory presentation was better than visual presentation. And, as before, the deaf children showed a reversal of this modality effect for learning. For these children, visual presentation was better than auditory presentation. Although none of the pair-wise differences in the sequence learning conditions reached statistical significance for the normal-hearing children, the overall pattern of their learning spans was similar to their earlier memory span results and to the pattern observed with the adults.

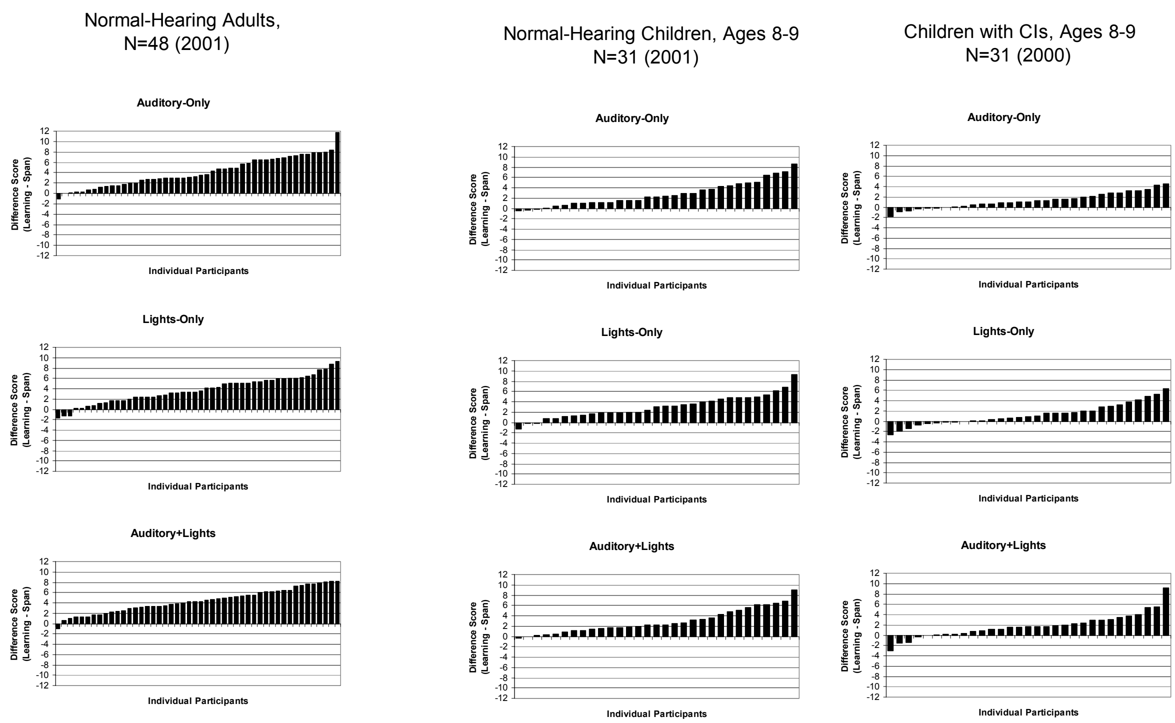


Figure 7. Difference scores for individual subjects showing sequence learning score minus his working memory span score. Data for the auditory-only (A) condition is shown on the top, lights-only (L) condition in the middle, and auditory-plus-lights (A+L) condition, on the bottom. Data from normal-hearing adults are shown on the left, scores for normal-hearing 8- and 9-year-old children in the center, and scores for 8- and 9-year-old cochlear implant users on the right.

To assess the magnitude of the repetition learning effects, we computed difference scores between the learning and memory conditions by subtracting the memory span scores from the learning span scores. The difference scores for the individual subjects in each group for the three presentation formats are displayed in Figure 7. Inspection of these distributions reveals a wide range of performance for all three groups of subjects. While most of the subjects in each group displayed some evidence of learning in terms of showing a positive repetition effect, there were a few subjects at the end of the distribution who either failed to show any learning at all or showed a small reversal of the predicted repetition effect. Although the number of these subjects was quite small in the adults and normal-hearing children, about one-third of the deaf children showed no repetition learning effect at all and no benefit from having the same stimulus sequence repeated on each trial.

Theoretical Significance

Taken together, the results of our recent experiments on working memory provide some new insights into the elementary information processing skills of deaf children with cochlear implants and the underlying cognitive factors that may affect their speech and language abilities. Our studies of working memory capacity using traditional digit span tests and the new Simon memory game were specifically designed to obtain process measures of performance that assessed specific subcomponents of working memory in order to understand the nature of the capacity limitations in encoding and processing sensory information. In this section, we briefly discuss the theoretical significance of our findings in light of the problems surrounding the enormous individual differences in the clinical outcome measures of speech and language that have been consistently reported in the literature.

Detailed analyses of the links between three different sets of measures—the digit span scores, the sentence durations and the word recognition scores using partial correlations, revealed that the common source of variance that was shared between all three of these tasks was processing speed, specifically, articulation speed and, by inference, speed of the verbal rehearsal process. This is a significant finding theoretically because it provides converging evidence from several different behavioral measures obtained on the same children for the existence and operation of a common information processing mechanism used for storage and maintenance of information in working memory and it suggests a principled explanation for the individual differences observed in a wide range of speech and language processing tasks. A task analysis of the traditional test battery used for assessment reveals that verbal rehearsal processes are a common component of every one of the outcome measures used to assess speech perception, spoken word recognition, vocabulary, comprehension and speech intelligibility.

We also found effects of early deafness and auditory deprivation on memory and learning of visual sequential patterns, a result that was initially unexpected when we began this research. The visual-only spans for the deaf children were shorter than the visual-only spans obtained from the age-matched normal-hearing children. This difference was observed in both the Simon memory span experiment, which used random sequences and the Simon learning experiment, which used repeated sequences.

The results of the visual-only conditions are of special theoretical interest to us because they demonstrate that differences observed in working memory are not necessarily restricted only to temporal patterns perceived via the auditory sensory modality; differences in memory span were also found in tasks when both item and order information in sequential visual patterns had to be preserved for a short period of time in immediate memory.

More detailed analyses of these results suggest that the deaf children may have used two different coding strategies to carry out the Simon sequence memory task, a visual-spatial coding strategy for the visual patterns and a verbal coding strategy for the auditory patterns. In contrast, the normal-hearing children may have used the same verbal coding and rehearsal strategy for both the visual and auditory patterns, a processing strategy that emerges, develops and becomes highly automatized over time and is routinely applied in a mandatory fashion for sequential patterns containing familiar linguistic stimuli like color names.

As we noted earlier in the section on redundancy gains under multi-modal presentation, the deaf children with cochlear implants did not use the informationally redundant auditory cues as well as the normal hearing children did to improve their immediate memory capacity. This finding was observed even when the deaf children could identify all of the auditory signals presented in isolation. To understand the basis for these differences and to obtain some further insights into the nature of the coding strategies used by the deaf children in carrying out the Simon memory game task, we first examined the intercorrelations between the multi-modal and the visual-only conditions of the Simon memory game for both groups of children (see also Cleary et al., 2001). We then examined the correlations between the WISC digit spans and the memory span scores obtained from the Simon memory game.

Table IV presents a summary of the intercorrelations among the two different sets of memory span scores, the forward and backward WISC digit span scores, and the Simon memory span scores for auditory, visual and auditory + visual presentation for both groups of children. The top section of this table displays the correlations among these memory span measures for the deaf children with cochlear implants, the bottom section displays the correlations for the normal-hearing children.

Examination of the intercorrelations of the Simon memory span scores for the deaf children in the top section of the table reveals a strong positive correlation between the multi-modal and the visual-only conditions ($r= +.71$, $p< .01$). If the deaf children were using the same visual-spatial coding strategy in both conditions of the Simon memory task, regardless of whether additional redundant auditory information were available, one would predict that these two measures of memory span would be strongly correlated. This is exactly the pattern we found. In contrast, examination of the intercorrelations for the normal-hearing children in the bottom section of the table reveals a different result. Here the correlation between the multi-modal and visual-only conditions is much lower and non-significant ($r= +.20$, NS). This pattern of results suggests that the normal-hearing children used two different coding strategies in the Simon memory game, a verbal-sequential strategy in the multi-modal condition and a visual-spatial strategy in the visual-only condition, while the deaf children used the same visual-spatial coding strategy in both tasks.

Additional support for this explanation comes from an analysis of the correlations between the WISC digit span scores and the Simon memory game scores. As shown in Table IV, a different pattern of correlations was observed between these two sets of measures in each group of children, suggesting again that different coding strategies were used to perform these two tasks. Although the WISC forward digit span scores were positively correlated with the memory game scores for the normal-hearing children, the correlation with forward span was only observed in the multi-modal Simon condition and not the visual-only Simon condition. This finding suggests that the normal-hearing children used the same verbal coding strategies in both the forward digit span task and the multi-modal Simon memory span task. In contrast, the WISC forward digit spans and the memory game conditions were not correlated at all for the deaf children in either of these conditions.

The dissociations observed between these two memory tasks would be expected if the deaf children used one coding strategy in the multi-modal condition of the Simon memory game and a different strategy in the WISC forward digit span task. The failure to find the same pattern of intercorrelations among the two different memory tasks in the deaf children with cochlear implants, taken together with the smaller redundancy gains observed in multi-modal Simon condition, suggests that the deaf children encoded and processed these temporal sequences in fundamentally different ways than the normal-hearing, typical-developing children. The deaf children apparently relied primarily on a visual-spatial coding strategy to perform the Simon memory game task even when additional redundant auditory cues were available to help them improve their performance on this memory task.

Table IV

Intercorrelations Obtained in Experiment 1 Between WISC Forward Digit Span, Memory Game Conditions, and WISC Backward Digit Span

Cochlear Implant Group, N = 43, (subgroup values in parentheses, N = 22)

	WISC Forward Digit Span	Memory Game: Color-names-plus-Lights (A+L)	Memory Game: Lights-only (L)
Memory Game Color-names-plus-Lights (A+L)	.09 (.14)		
Lights-Only (L)	-.02 (-.07)	.71** (.60**)	---
WISC Backwards Digit Span	.52** (.47*)	.46** (.50*)	.48** (.57**)

Normal-hearing Group, N = 44, (subgroup values in parentheses, N = 22)

	WISC Forward Digit Span	Memory Game: Color-names-plus-Lights (A+L)	Memory Game: Lights-only (L)
Memory Game Color-names-plus-Lights (A+L)	.58** (.54*)	---	
Lights-Only (L)	.01 (.03)	.20 (.30)	---
WISC Backwards Digit Span	.32* (.08)	.36* (.52*)	.19 (.22)

Note: *$p < .05$. **$p < .01$. Although age in months has been partialled out of the calculations, age showed only a small positive correlation in the NH group and a near-zero correlation in the CI group, making the simple correlations almost identical to those shown above.

Final Remarks on Process Measures of Performance

The new findings summarized in this chapter suggest that additional processing measures of performance should be developed to study other aspects of cognition, such as attention, categorization, learning and memory—“central” cognitive processes that make use of the initial sensory input provided by a cochlear implant. One can imagine the construction of an entirely new battery of behavioral tests based on “process” measures of performance that could be used to assess the benefits of cochlear implantation and to study the time course of development of these basic information-processing skills. Some of these measures could be used to assess how well a listener is able to use the limited and impoverished sensory information conveyed through the cochlear implant. Other measures could be used to assess differences in processing speed, efficiency and processing capacity. Additional measures of working memory span, verbal and visual-spatial coding and rehearsal strategies, controlled and automatic attention and the development of automaticity may provide new knowledge about individual differences in the elementary cognitive processes that underlie the traditional endpoint measures of outcome performance. At some point in the future, it may also be possible to develop a set of pre-implant, performance-based measures of visual attention and memory for temporal sequences that could be used to predict outcome and benefit after implantation. Given the new tools and experimental methodologies that are currently available from research in cognitive science, we believe these goals can be achieved in the next few years as the age of implantation becomes lower and more profoundly deaf children become candidates for cochlear implantation.

References

Baddeley, A., Gathercole, S.E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A.D., Thomson, N. & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575-589.

Ball, G.F. & Hulse, S.H. (1998). Birdsong. *American Psychologist, 53,* 37-58.

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.

Biederman, I. & Shiffrar, M.M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13,* 640-645.

Carpenter, P.A., Miyake, A., & Just, M.A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging, In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1075-1122). San Diego, CA: Academic Press.

Cleary, M., Pisoni, D.B., & Geers, A.E. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear & Hearing, 22,* 395-411.

Cowan, N., Wood, N.L., Wood, P.K., Keller, T.A., Nugent, L.D. & Keller, C.V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*, 141-160.

Cullington, H., Hodges, A.V., Butts, S.L., Dolan-Ash, S. & Balkany, T.J. (2000). Comparison of language ability in children with cochlear implants placed in oral and total communication educations settings. *Annals of Otology, Rhinology and Laryngology, 185*, 121-123.

Ericsson, K.A. & Pennington, N. (1993). The structure of memory performance in experts: Implications for memory in everyday life. In G.M. Davies & R.H. Logie (eds.), *Memory in everyday life*, pp. 241-282. Elsevier.

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). St. Louis, MO: Central Institute for the Deaf.

Gupta, P. & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain & Language*, 59, 267-333.

Haber, R.N. (1969). *Information-processing approaches to visual perception.* New York: Holt, Rinehart and Winston.

Hodges, A.V., Dolan-Ash, M., Balkany, T.J, Schloffman, J.J. & Butts, S.L. (1999). Speech perception results in children with cochlear implants: contributing factors. *Otolaryngology-Head & Neck Surgery, 12*, 31-34.

Kirk, K.I., Diefendorf, A.O., Pisoni, D.B. & Robbins, A.M. (1997). Assessing speech perception in children. In L.L. Mendel and J.L. Danhauer (Eds.), *Speech Perception Assessment*, Pp. 101-132. San Diego: Singular Press.

Kirk, K.I., Miyamoto, R.T., Lento, C., Ying, E., O'Neill, T. & Fears, B. (in press). Effect of age at implantation in young children. *Annals of Otology, Rhinology, and Laryngology*.
Kirk, K.I., Pisoni, D.B., & Miyamoto, R.T. (2000). Lexical discrimination by children with cochlear implants: Effects of age at implantation and communication mode. In S.B. Waltzman & N.L. Cohen (Eds.), *Cochlear Implants* (pp. 252-254). New York: Thieme.
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.
Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience, 8,* 125-170.
Konishi, M., & Nottebohm, R. (1969). Experimental studies in the ontogeny of avian vocalizations. In R.A. Hinde (Ed.), *Bird Vocalizations* (pp. 29-48). New York: Cambridge University Press.
Lachman, R., Lachman, J.L., & Butterfield, E.C. (1979). *Cognitive psychology and information processing: An introduction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.
Levitt, H. (1970). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49,* 467-477.
Lindsay, P.H., & Norman, D.A. (1977). *Human information processing : An introduction to psychology, 2nd edition.* New York: Academic Press.
Luce, P.A., & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing, 19,* 1-36.
Marler, P., & Peters, S. (1988). Sensitive periods for song acquisition from tape recordings and live tutors in the swamp sparrow, *Melospiza georgiana*. *Ethology, 77,* 76-84.
McGarr, N.S. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research,* 26, 451-458.
Meyer, T.A. & Pisoni, D.B. (1999). Some computational analyses of the PBK Test: Effects of frequency and lexical density on spoken word recognition. *Ear & Hearing, 20,* 363-371.
Miyamoto, R.T., Svirsky, M.A. & Robbins, A.M. (1997). Enhancement of expressive language in prelingually deaf children with cochlear implants. *Acta Otolaryngol, 117,* 154-157.
NIH Consensus Conference. (1995). Cochlear implants in adults and children. *Journal of the American Medical Association, 274,* 1955-1961.
Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.
Penny, C.G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*, 398-422.
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., Geers, A.E. & Tobey, E.A. (2000). Individual differences in effectiveness of cochlear implants in prelingually deaf children: Some new process measures of performance. *Volta Review, 101,* 111-164.
Pisoni, D.B. & Geers, A. (2000). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing. *Annals of Otology, Rhinology & Laryngology, 109*, 92-93.
Pisoni, D.B., Svirsky, M.A., Kirk, K.I., & Miyamoto, R.T. (1997). Looking at the "Stars": A first report on the intercorrelations among measures of speech perception, intelligibility, and language development in pediatric cochlear implant users. *Research on Spoken Language Processing Progress Report No. 21*, (pp. 51-91). Bloomington, IN: Speech Research Laboratory.
Reitman, W.R. (1965). *Cognition and thought: An information-processing approach.* New York: Wiley.
Robbins, A. & Kirk, K.I. (1996). Speech perception assessment and performance in pediatric cochlear implant users. *Seminars in Hearing*, 17, 353-369.
Robbins, A.M., Svirsky, M., Osberger, M.J. & Pisoni, D.B. (1998). Beyond the audiogram: The role of functional assessments. In F. Bess (Ed.), *Children with Hearing Impairment*. Pp. 105-124. Nashville, TN: Bill Wilkerson Center Press.
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.

Rudel, R.G. & Denckla, M.B. (1974). Relation of forward and backward digit repetition to neurological impairment in children with learning disability. *Neuropsychologia, 12*, 109-118.

Shadmehr, R., & Holcomb, H.H. (1997). Neural correlates of motor memory and consolidation. *Science, 277,* 821-825.

Svirsky, M.A., Robbins, A.M., Kirk, K.I., Pisoni, D.B. & Miyamoto, R.T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science, 11,* 153-158.

Svirsky, M.A., Sloan, R.B., Caldwell, M. & Miyamoto, R.T. (2000). Speech intelligibility of prelingually deaf children with multichannel cochlear implants. *Annals of Otology, Rhinology, and Laryngology, 185,* 123-125.

Tait, M., Lutman, M.E. & Robinson, K. (2000). Preimplant measures of preverbal communicative behavior as predictors of cochlear implant outcomes in children. *Ear & Hearing, 21*, 18-24.

Tobey, E.A., Geers. A.E., Morchower, B., Perrin, J., Skellett, R., Brenner, C., & Torretta, G. (2000). Factors associated with speech intelligibility in children with cochlear implants. *Annals of Otology, Rhinology and Laryngology Supplement, 185*, 28-30.

Watkins, M.J., Watkins, O.C. & Crowder, R.G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior, 13*, 430-447.

Wechsler. D. (1991). *Wechsler Intelligence Scale for Children, Third Edition (WISC-III).* San Antonio, TX: The Psychological Corporation.

Zwolan, T.A., Zimmerman-Phillips, S., Asbaugh, C.J., Hieber, S.J, Kileny, P.R. & Telian, S.A. (1997). Cochlear implantation of children with minimal open-set speech recognition skills. *Ear & Hearing*, 18, 240-251.

Word-Learning Skills of Profoundly Deaf Children Following Cochlear Implantation: A First Report[1]

Derek M. Houston,[2] Allyson K. Carter, David B. Pisoni,[2] Karen Iler Kirk,[2] and Elizabeth A. Ying[2]

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

[1] This research was supported by NIH Research Grants DC00111, DC00064 and NIH T32 Training Grant DC00012 from the NIDCD to Indiana University. We are grateful to Sister Mary Gannon and her staff at St. Joseph's Institute for the Deaf, St. Louis, MO, Liane Garner and her staff at St. Joseph's Institute for Deaf, Kansas City, KS, Michele Wilkins and her staff at Child's Voice, Elmhurst, IL, Maria Sentelik and her staff at Ohio Valley Oral School, Cincinnati, OH, and Beth Jeghlum and her staff at the Center for Young Children daycare center, Indianapolis, IN for their invaluable help in recruiting and providing testing places for the children in this project. We also wish to thank Miranda Cleary, Caitlin Dillon, Cara Lento, Tara O'Neill, Helen Zuganelis, and Amy Perdew for their help with data collection.
[2] Also DeVault Otologic Research Laboratory, Department of Otolaryngology–Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, Indiana.

Word-Learning Skills of Profoundly Deaf Children Following Cochlear Implantation: A First Report

Abstract. In recent years, cochlear implant (CI) technology has advanced substantially. Deaf children can now be provided with an electrical signal that codes sound input to facilitate spoken language learning. However, a great deal of variability has been observed in the audiological outcome measures obtained from pediatric CI recipients. Deaf children who have received CIs often lag substantially behind their normal hearing (NH) peers on a wide range of speech and language measures. Many factors contribute to this variability in performance, including age of implantation, amount of speech therapy, cognitive information processing factors, such as working memory span, as well as numerous linguistic factors. An important fundamental linguistic skill that plays a critical role in later language development is novel word learning, the ability to map the sound patterns of spoken words onto their physical referents. This paper describes an experimental procedure that was developed to investigate the word-learning skills of deaf children who have received CIs and reports some preliminary findings obtained from 2- to 5-year-old CI users. Each child was presented with a set of Beanie Babies™ and a set of associated labels for their names using interactive play scenarios. After training, the children's receptive and expressive knowledge of the new names was tested both immediately after exposure and then following a 2-hour delay. Pediatric CI users performed more poorly than age-matched NH children on both receptive and expressive word-learning tests. Both groups of children showed retention of the new names from immediate to delayed testing conditions for words that were previously known by the children. However, additional analyses of the deaf children's response revealed that they performed more poorly after a delay on both receptive and expressive tests when the words were unfamiliar to them before coming into the laboratory. Our findings on novel word learning suggest that although pediatric CI users may have impaired and/or delayed phonological processing skills their long-term memory for familiar spoken words that they are able to perceive and encode appears to be similar to NH children. Implications of these findings for receptive and expressive language development in this clinical population are discussed.

Introduction

Cochlear implants (CIs) provide profoundly deaf individuals with access to sound and offer the possibility of learning spoken language by exposure to an electrical signal that codes the auditory input. For postlingually deafened adults, CIs are a way of restoring hearing. For profoundly deaf children, CIs represent a new sensory modality that provides them with an opportunity to hear speech in their environment and learn spoken language. Several investigations have reported that cochlear implantation improves deaf children's language comprehension and production skills (Eisenberg, Martinez, Sennaroglu, & Osberger, 2000; Kirk, Osberger, Robbins, Riley, & Todd, 1995; Osberger, Robbins, Todd, & Riley, 1994; Tobey, Geers, & Brenner, 1994; Tomblin, Spencer, Flock, Tyler, & Gantz, 1999; Zwolan et al., 1997). However, these investigations also have revealed that deaf children with CIs do not perform as well as their normal-hearing (NH) peers on a wide range of speech and language measures and that there are enormous individual differences in language skills after implantation (Pisoni, Cleary, Geers, & Tobey, 2000). In one study, Svirsky, Robbins, Kirk, Pisoni, and Miyamoto (2000) followed deaf children's language development before and at regular 6-month intervals after cochlear implantation for three years. They found that while deaf children performed better on language measures after cochlear implantation than would be predicted using conventional hearing aids, there was considerable individual variability among the deaf children, and their language skills continued to lag behind their NH peers. Other studies have reported that the vocabulary levels of deaf and hard-of-hearing children lag behind those of NH children (Lederberg & Spencer, 2001).

Several factors may contribute to the generally poorer performance on spoken language measures of deaf children who use CIs compared to their NH peers. The electrical signal transmitted by a CI is highly degraded compared to a normal acoustic signal. Thus, the initial quality of the speech signal that a deaf child hears is much poorer than the speech signals received by NH children. Also, congenitally deaf children who receive CIs have had some period of auditory deprivation caused by deafness, so they are exposed to less auditory input during critical stages of speech perception and language development than NH children.

Normal-hearing children also develop spoken language at much faster rates than deaf children who receive CIs. However, there are enormous individual differences in language performance among deaf children who use CIs. Some of these children acquire language quite well while others continue to be delayed relative to their NH peers (Pisoni et al., 2000). The primary demographic factors that have been associated with differences in audiological outcome measures among these children are duration of deafness, length of CI use, and age at implantation (Kirk, 2000). Several other factors also affect performance. These include the type and amount of therapy, etiology of hearing loss, number of electrodes inserted into the cochlea (Loizou, Dorman, & Tu, 1999), as well as the child's early linguistic environment and socioeconomic factors (Hart & Risley, 1995).

Another contributing source of variability is related to differences in underlying cognitive processing skills. Recently, Pisoni, Svirsky, Kirk, and Miyamoto (1997) reported that CI children's performance on speech perception, speech production, and language tests was highly correlated (see also Pisoni et al., 2000). The authors suggested that the common variance observed in these tasks might be attributed to some underlying skills, including the phonological encoding, storage, and retrieval of spoken words. It is very likely that auditory deprivation experienced by the congenitally deaf during early childhood may lead to delays and/or disorders in these kinds of cognitive processing skills. Moreover, loss of these skills may be responsible for the poorer linguistic performance observed in pediatric CI users relative to NH children.

A number of basic cognitive and linguistic skills may become delayed and/or disordered in this clinical population and affect their spoken language development. One of the building blocks of language development is the ability to learn words. Word learning requires listeners to encode the sound patterns of words into memory and associate these sound patterns with objects, actions, and concepts in the world. The sound-meaning associations that are learned must be maintained in memory and remain accessible for future use. Thus, phonological processing skills and memory are crucial for children's word learning and ultimately their language development. A period of auditory deprivation during development may have effects on how auditory information is encoded and maintained in memory even after interventions such as cochlear implantation. Thus, congenitally deaf children who have virtually no access to speech sounds until intervention may have particular difficulty with linguistic skills that involve perceiving and interpreting speech sounds and learning new words. The aim of this project was to develop procedures to investigate the early word-learning skills of deaf children following cochlear implantation and explore differences in performance between deaf children and NH children.

Normal-hearing children begin producing words at approximately 12 months of age. By 18 months, most infants can produce over 50 words and begin learning several words each day (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994). Most research on word learning focuses on how children learn to correctly associate the sound patterns of words with their referents (e.g., Clark, 1973, 1983;

Markman, 1991; Nelson, 1988). Recently, some work has explored young children's ability to encode the sound patterns of novel words. Jusczyk and colleagues have shown that by 8 months of age, infants can encode the sound patterns of words and store them in memory (Houston & Jusczyk, submitted; Jusczyk & Aslin, 1995; Jusczyk & Hohne, 1997). The ability to encode phonological information into memory enables children to form lexical representations and eventually build a lexicon of the words in the target language.

Phonological encoding has been shown to play an important role in vocabulary development. Gathercole and Baddeley (1989, 1990) have found a strong relationship between phonological working memory span and vocabulary size in NH children. They suggest that proficient word learners are able to maintain words in working memory by internal verbal rehearsal procedures that facilitate the encoding of words into long-term memory and development of lexical representations. Repetition is also important for encoding words into memory. Huttenlocher, Haight, Bryk, Seltzer, and Lyons (1991) found a significant correlation between how often parents use words and children's acquisition of those words, suggesting that frequency of exposure affects how quickly children learn words (see also Hart & Risley, 1995). Taken together, these findings suggest that encoding phonological representations of words and maintaining them in memory are important underlying skills that affect subsequent novel word-learning and language development.

To build a vocabulary, children not only have to encode the phonological representations of spoken words but they must also map phonological representations onto meanings. In the late 1970s, Susan Carey and her colleagues suggested that learning the meaning of words involves two stages (Carey, 1978; Carey & Bartlett, 1978). The first stage, which they called "fast mapping," involves the initial encoding of the sound pattern of the words and some basic understanding of the meaning. The second stage involves developing a fuller, more detailed understanding of words by hearing them in several different contexts so that hypotheses about their meaning can be tested. In a well-known study, Carey and Bartlett (1978) used a novel word in a casual undirected way to NH preschool children (e.g., "…bring me the *chromium* one. Not the red one…") in the presence of an olive and a red colored tray. They found that after only a single presentation of the novel word, the children began forming basic hypotheses about the meaning of the word. For example, when asked later by the experimenter, some of the children indicated that the word chromium was a color.

In another study of fast mapping, Heibeck and Markman (1987) reported that children as young as two years of age showed fast mapping of shape and texture terms as well as color terms. More recently, Markson and Bloom (1997) tested the ability of NH 3- and 4-year-olds to learn novel names for objects and remember them over a delay. Children participated in several tasks with the experimenter where they played with 10 objects and were casually taught a novel name (e.g., "koba") for one of the novel objects. Although they were not asked to repeat the name or even acknowledge that they had heard it. The children were then tested in one of three delay conditions (immediate, 1 week, 1 month) in which they were presented with the original 10 objects and asked to identify the object called "koba." Markson and Bloom found that children performed well above chance in selecting the correct object in all delay conditions. There was also no main effect of delay, suggesting that when NH children learn novel words, lexical representations persist in long-term memory. While a complete understanding of novel words may involve a complex and lengthy process, the basic process appears to begin with an initial "fast mapping" stage of word learning that is immediate and obligatory in establishing a solid foundation for later lexical development.

The fast mapping stage of early word learning requires children to encode the phonological form of spoken words very rapidly on-the-fly without repetition or practice. Children who have difficulty with speech perception and phonological encoding may show difficulties in learning novel words. There may be a high incidence of poor phonological encoding skills among CI users for several reasons. First, as noted earlier, the auditory information provided by a CI is impoverished and highly degraded when compared to normal hearing. The cochlea of a NH adult has approximately 13,500 outer hair cells and 3500 inner hair cells that respond to different acoustic frequency ranges and contribute to stimulating the spiral ganglion cells. In contrast, cochlear implants bypass these hair cells and other cochlear mechanisms to stimulate the spiral ganglion cells with only a few dozen electrodes. Not surprisingly, the frequency resolution of sound provided by a cochlear implant is not as high as what a healthy cochlea normally provides. As a consequence, it is possible that the initial sensory input from a cochlear implant may be a limiting factor in encoding the sound pattern of spoken words.

Another reason why pediatric CI users may not develop word-learning skills as well as NH children is that the period of early sensory deprivation prior to implantation may lead to a delayed and/or disordered course of language development. There is some evidence to suggest that any degree of hearing loss may cause problems in phonological processing and word learning. For example, Gilbertson and Kamhi (1995) assessed hearing-impaired children's ability to encode phonological information and learn novel words while wearing their hearing aids. They found that hearing-impaired children's unaided level of hearing loss (ranging from mild to moderate) did not correlate significantly with word-learning abilities. However, the ability to encode phonological information did correlate with word-learning proficiency. Gilbertson and Kamhi concluded that even a mild hearing loss was a significant risk factor for language impairments related to the development of phonological processing skills. One possible reason the investigators did not find that unaided degree of hearing loss was an important factor in word-learning could be that the children's hearing may have been improved enough to learn words from incomplete or partial phonological representations when using hearing aids. It is possible that pediatric CI users who have more substantial profound hearing losses may display similar or even more severe long-term phonological processing deficits because, prior to CI intervention, they experienced long periods of auditory deprivation that could not be ameliorated by conventional hearing aids.

Little is currently known about the word-learning skills of prelingually deaf children who have a severe to profound hearing impairment. Recent work by Lederberg and colleagues has begun to explore the vocabulary development of deaf and hard-of-hearing children who use both spoken and sign language. In one study, Lederberg, Prezbindowski, and Spencer (2000) tested deaf and hard-of-hearing children's ability to learn novel object nouns in several contexts in which they were given either implicit or explicit reference to the novel objects with which they were to associate the novel words. They also assessed the children's vocabulary using the MacArthur Communication Development Inventory (CDI). The CDIs are parent report forms that are used to assess individual infants' and young children's language and communication skills (Fenson et al., 1994). The forms consist of lists of gestures, words, and sentences, and parents are asked to indicate whether or not their child understands and/or produces any of the items.

Lederberg and colleagues found that deaf and hard-of-hearing children who have larger vocabularies in their preferred mode of communication (i.e., spoken language vs. sign) were able to learn words using their preferred mode of communication with fewer exposures and with less explicit teaching than children with relatively smaller vocabularies. They concluded that there is a relationship between vocabulary size and word encoding skills: a larger vocabulary appears to be related to learning words by implicit reference. Lederberg et al.'s study is an important step in understanding the word-learning skills and vocabulary development in deaf children who use manual and/or oral communication. However, there are no studies in the literature that focus specifically on the spoken word-learning skills of profoundly deaf children who received CIs and use oral-aural communication.

The acquisition of word-learning skills by deaf and hard-of-hearing children with CIs might take on a very different developmental course depending on whether the children use manual or oral-aural communication methods. Deaf and hard-of-hearing children may be able to learn words in the manual communication mode as readily as NH children do in the oral communication mode. However, deaf CI users' ability to encode spoken words aurally and become oral communicators may be quite different because they receive degraded and highly impoverished acoustic representations of speech that encode partial information about the phonological structure of the sound patterns of the language. Deaf and hard-of-hearing children have poorer phonological processing skills than NH children (Briscoe, Bishop, & Norbury, 2001; Gilbertson & Kamhi, 1995) and poorer phonological processing skills have been shown to be related to poorer performance on oral word-learning tasks (Gilbertson & Kamhi, 1995). Assessing the early word-learning skills of deaf children who are oral-aural communicators and who use cochlear implants is an important step for understanding the linguistic processes that play a role in deaf children's phonological processing skills and learning spoken language after received a CI. The time course of deaf children's language development may be delayed and/or disordered compared to NH children. Moreover, it is possible that phonological processing disorders may contribute to the enormous individual variability observed/reported among pediatric CI users on a wide range of speech and language outcome measures.

To date, we know of no other investigations of the word-learning skills of deaf children following cochlear implantation. This particular group of children is an unusual and unique clinical population because until the time they receive their CIs, their linguistic, cognitive, and perceptual skills have developed in the absence of sound and input from the auditory sensory modality. Once these children receive their CIs, they are able to gain access to auditory information and begin to acquire knowledge of speech and properties of spoken language. The initial goal of this study was to develop a set of experimental procedures that could be used to determine whether CIs could enable deaf children to develop the skills necessary to learn spoken words rapidly after only a few exposures. We were interested in exploring how quickly deaf children with cochlear implants can map sound patterns onto referents in a laboratory setting. We were also interested in determining the relationship between measures of early word learning and other audiological outcome measures such as spoken word recognition, speech intelligibility and receptive and expressive language abilities. Another goal of this study was to assess pediatric CI users' long-term memory for spoken words. While NH children may display little difficulty learning novel words and recalling them after a short period of time, deaf children with CIs may not have developed the same memory skills. It is possible that early auditory deprivation may not only affect children's immediate phonological processing skills, but also their ability to encode, verbally rehearse and store the sound patterns of spoken words in long-term memory.

This report describes the results of a preliminary study that assessed the abilities of preschool-aged NH and hearing-impaired children with CIs to learn words quickly after a brief exposure period. Specifically, we developed a name-learning task using a set of Beanie Baby™ stuffed animals. The first part of the project involved selecting new names for the Beanie Baby™ animals that would be taught to the children. This was done by eliciting names from adult participants that corresponded to perceptual attributes of the animals. In the second part of the project, we conducted a novel word-learning experiment with these materials. All children were taught the names for a set of Beanie Babies™, one-at-a-time, and were then tested receptively and expressively for their knowledge of the names both immediately after learning and then following a 2-hour delay. We report data from a group of 24 deaf children who use CIs and a group of 24 age-matched NH children who served as a comparison group.

Construction of Stimulus Materials

The stimulus materials used in the experiment reported below consisted of a set of sixteen Beanie Baby™ stuffed animals. Each Beanie Baby™ comes with a name assigned by the manufacturer (Ty Corporation®). We did not use these names because some children might already know them while others might not and because some of the names were related to physical attributes of the stuffed animals while others were not. We obtained a set of new names from a group of adult participants that corresponded to salient physical attributes of the Beanie Babies™. This was done to facilitate learning an association between the attribute names and the referent Beanie Babies™. Because many of the Beanie Babies™ have several perceptual features that could be considered perceptually salient, a pilot experiment was conducted to determine which characteristics of each Beanie Baby™ were most distinctive. The goal of this initial study was to select a label for each Beanie Baby™ that could then be used in the subsequent word-learning experiment with children.

Method

Participants. The participants were 37 undergraduates at Indiana University who reported no prior history of speech or hearing disorders. Thirty-five of the subjects were native English speakers. All subjects were recruited from the Indiana University community and all received partial credit towards an Introductory Psychology class requirement for their participation. The mean age of the participants was 19.9 years (SD = 1.3).

Materials. Sixteen Beanie Baby™ stuffed animals were used as stimuli. These Beanie Babies™ were selected from a larger set of Beanie Babies™ on the basis of whether they had distinguishing perceptual attributes that could be easily named, such as a long tail, horns, or a bright color.

Procedure. Subjects were tested in three groups in a small experimental classroom. They were given written instructions in which they were told that the experimenter would hold up each of the sixteen Beanie Babies™ individually, and they would be asked to invent new names for the Beanie Babies™, as if they were teaching the names to a young child. Subjects were asked to re-name the original Beanie Babies™ using names that described some physical attributes of the Beanie Babies™. Subjects were instructed to provide up to three new names for each animal, and to use one-word names only. Subjects were provided with response sheets on which to write the new names. The Beanie Babies™ were presented individually, one at a time, in a random order to the three groups.

Results

For each Beanie Baby™, the responses were recorded and tallied to calculate the frequency of the names generated. A new Beanie Baby™ attribute name was chosen from the response distributions based on two criteria: (1) that the name was the most frequent response among students, and (2) that it reflected a true physical attribute of the animal. For example, the name “Red” was the most frequent response and was also an appropriate name for the red bull because it refers to the color of the bull. In contrast, a non-attribute name, “Teddy,” was the most frequent response for the brown bear, but was inappropriate for our purposes because it did not represent a physical attribute. The second most common response was “Fuzzy”, which we used since it describes a perceptual attribute of the bear. The new attribute name was the most frequent response for seven of the Beanie Babies™ (“Blue,” “Red,” “Stripes,” “Pink,” “Spots,” “Ears,” “Tail”), the second most frequent response for four of the animals (“Wings,” “Fuzzy,” “Legs,” “Cottontail”), the third most frequent response for five of the animals (“Horns,” “Gray,” “Teeth,” “Bushy”), and the fourth most frequent response for “White.” Table 1 provides a list of each of the original Beanie Baby™ stuffed animal names and descriptions, the new attribute names derived from this procedure, along with the percentage of subjects who used the new attribute name.

Original Beanie Baby™ name	"New" attribute name	Frequency of new name response (%)	Original Beanie Baby™ name	"New" attribute name	Frequency of new name response (%)
Crunch the Shark	*Teeth*	11.8	*Dotty* the Dalmatian	*Spots*	46.6
Rocket the Bird	*Blue*	36.6	*Halo* the Angel Bear	*White*	6.1
Batty the Bat	*Wings*	13.7	*Spunky* the Cocker Spaniel	*Ears*	14.9
Kuku the Bird	*Pink*	29.6	*Nuts* the Squirrel	*Bushy*	11.1
Snort the Bull	*Red*	32.4	*Nibbly* the Bunny	*Cottontail*	14.1
Goatee the Goat	*Horns*	8.2	*Spinner* the Spider	*Legs*	14.9
Buster the Bear	*Fuzzy*	9.8	*Prance* the Tabby Cat	*Stripes*	25.6
Tiptoe the Mouse	*Tail*	18.5	*Spike* the Rhino	*Gray*	7.8

Table 1. Original names of Beanie Babies™, new given attribute names, and the frequency with which each new name was generated.

Novel Word-learning Experiment

In order to establish a starting point for future research, we designed a task that we thought would optimize the children's performance in learning the novel names of Beanie Babies™. To study early word-learning skills in these children, we asked the children to learn novel names of Beanie Babies™ that corresponded to salient attributes, although in most cases each attribute name fit more than one Beanie Baby™. Two groups of children were studied. One group was coded as "young" and included children who were 2 and 3 years of age, and one group was coded as "old" and included children who were 4 and 5 years of age. The young children were taught the names of eight Beanie Babies™ (in two sets of four) while the old children were taught the names of sixteen (in two sets of eight). After training, children were tested for their receptive knowledge of the names using a forced-choice recognition task. They were also tested for their expressive knowledge of the names, using a cued-recall task. Finally, their long-term memory of the names was subsequently assessed in a second session by re-testing them at least two hours later on both receptive and expressive tasks. We predicted that if CI children have impaired phonological processing and word-learning skills, they should not perform as well as NH children on the receptive and expressive tasks. Furthermore, if pediatric CI users' ability to maintain spoken words in memory is impaired, they might also show a greater effect of delay on performance than NH children. Finally, we predicted that if word-learning skills are an important source of variance in linguistic performance, then deaf children who use CIs should display a wider range of individual variability in their performance on these tasks than NH children.

Method

Participants. *CI Children:* Twenty-four prelingually deaf children who use cochlear implants were recruited from five locations: Indiana University School of Medicine, Indianapolis, Indiana; St. Joseph's Institute for the Deaf, St. Louis, Missouri; St. Joseph's Institute for the Deaf, Kansas City, Kansas; Child's Voice, Chicago, Illinois; and Ohio Valley Oral School, Cincinnati, Ohio. Three criteria were used for inclusion in this study. First, at the time of testing the deaf children had to be between the ages of 2;0 and 5;11. Second, the children had to be oral-only communicators (i.e., not use any form of manual communication). Third, the children had to have had at least one year of experience with their CI. The twelve children who fell into the young group were two and three years of age (mean = 37.7, and the remaining twelve children who fell into the old group were four and five years of age.[3] Table 2 displays the children's age at testing, age at which they received their CI, and the length of CI use.

NH Children: Twenty-four age-matched normal-hearing controls were recruited from the Bloomington, Indiana area and the Center for Young Children daycare center on the Indiana-University-Purdue-University-Indianapolis campus. Twelve children were assigned to the young NH group and twelve were assigned to old NH children group.[4] Ages are displayed in Table 2.

Group	Age at Testing		Age at Implantation		Length of CI use	
	Mean	Range	Mean	Range	Mean	Range
Young CI	37.7	28 – 46	20.6	10 – 34	17.1	12 – 26
Old CI	59.5	49 – 68	37.3	25 – 55	22.3	12 – 37
Young NH	28.0	28 – 46				
Old NH	58.4	49 – 64				

Table 2. Chronological age at testing, age at initial stimulation of cochlear implant, and length of cochlear implant use for the young and the old groups of pediatric cochlear implant users. Ages are given in months.

Materials. The stimulus materials consisted of 16 Beanie Babies™ that were assigned names by normal-hearing college students (see Stimuli Selection above). Each name corresponded to a salient physical attribute (e.g., "Red" is a red bull). The Beanie Babies™ were grouped into sets of four as shown in Table 3. The Beanie Babies™ were selected so that most of the attribute names could describe at least two Beanie Babies™ in the group. For example, "Wings," "Pink," and "Blue" all have wings. This was done so that the children would not be able to completely rely on a unique perceptual feature to identify the attributes in the tasks. For example, when they were asked to identify "Wings," three of the four Beanie Babies™ had wings, so they had to learn to map a specific feature to a specific Beanie Baby™.

Set	Beanie Baby™ Attribute Name	Description
A	Teeth	Shark
	Blue	Blue jay
	Wings	Bat
	Pink	Cockatoo
B	Red	Bull
	Horns	Goat
	Fuzzy	Brown bear
	Tail	Rat
C	Spots	Dalmatian
	White	White bear with halo
	Ears	Cocker spaniel
	Bushy	Squirrel
D	Cotton tail	Rabbit
	Legs	Spider
	Stripes	Cat with stripes
	Gray	Rhino

Table 3. Word set stimuli.

[3] The CI children's demographics were as follows: 1 Hispanic or Latino (male), 1 Asian/Pacific Islander (male), 2 Black/African American (1 male, 1 female), 20 White/Non Hispanic (14 male, 6 female).
[4] The NH children's demographics were as follows: 1 Hispanic or Latino (female), 1 Asian/Pacific Islander (male), 20 White/Non Hispanic (9 male, 11 female), 2 Other/Unknown (2 male).

Procedure. Children were explicitly taught a set of names for four or eight Beanie Babies™ using play scenarios (Training Phase 1). They were then given receptive (forced-choice) and expressive (cued-recall) tests for that set (Testing Phase 1). Children were then taught another set of Beanie Babies™ (Training Phase 2) and subsequently given the same tests with the second set (Testing Phase 2). Finally, after at least a two-hour delay[5], the children were given the same receptive and expressive tests over again, first with the Beanie Babies™ presented in Training Phase 1 and then with the set presented in Training Phase 2.[6]

Training Phase 1. *Young Children.* Each child was exposed to a set of four Beanie Babies™. Before the experiment, the exact order of Beanie Baby™ presentation was randomized and recorded on a form that was then followed during the experiment for each child. One experimenter (Experimenter 1) interacted with the child while a second experimenter (Experimenter 2) assisted the first experimenter in following the correct presentation order. The second experimenter also recorded the children's responses and tallied the number of times Experimenter 1 produced the name of each Beanie Baby™.

Experimenter 1 presented each Beanie Baby™, one at a time, to the child in a play scenario. Several toy props were used to create different play scenarios with each Beanie Baby™ in order to keep the task interesting. During each play interaction with a Beanie Baby™, Experimenter 1 used the name of the Beanie Baby™ exactly eight times. During the play scenario, Experimenter 1 tried to elicit three productions of the name from the child by encouraging the child to repeat the name after the Experimenter. Experimenter 2 recorded how many times the child produced each name. Positive feedback was given when the child produced the correct names. See Appendix for a sample scenario.

Older Children. The training phase was the same with the older children as with the younger children, except that eight Beanie Babies™ (two sets from Table 3) were taught during each phase instead of only four.

Testing Phase 1. The Testing Phase used two procedures to assess learning – a receptive test, which employed a forced-choice identification task, and an expressive test, which used a cued-recall task. Both tasks were given immediately after the training phase was completed Children were not given feedback about whether or not they gave correct responses during either of the test phases.

Receptive Test. For the forced-choice receptive test, all of the Beanie Babies™ (four or eight) were placed in a row in front of a child but they were initially hidden from view with a piece of cardboard. Then a toy bus or truck was brought out and placed in front of the child. Experimenter 1 then asked the child to "please put {*one of Beanie Babies™*} into the truck {*or bus*}." The child was encouraged to select one of the Beanie Babies™ in front of him/her but was not given any feedback as to whether the response was the correct choice. For example, the experimenter said "thank you," "good job," or clapped when the child made a selection, regardless of whether or not the correct Beanie Baby™ was selected. The Beanie Baby™ was then placed back in the row and the next trial was initiated. Each Beanie Baby™ was requested exactly once during each test phase.

Expressive Test. For the expressive test, Experimenter 1 played a "knock knock" game with the child. One Beanie Baby™ was placed behind a toy doorway. Experimenter 1 and/or the child said,

---

[5] For sixteen of the normal hearing children, the Long-term Memory test was conducted within 30 minutes after Testing Phase 2 (or approximately 40 to 60 minutes after the beginning of Training Phase 1). However, on all four of the tests, their performance did not differ significantly from those who had the full 2-hour delay. Thus, we included their data in our analyses.

[6] Two of the young CI children and one of the old NH children were not able to participate in the long-term memory portion of the experiment due to scheduling conflicts. In the analyses, missing cells were filled by computing the means of related data (i.e. scores on the same test by participants from the same subject group).

"knock knock," the door would open, and Experimenter 1 would ask the child, "Who's there?" The child was asked to name the Beanie Baby™, up to three times, until the child gave a response. Experimenter 2 recorded the child's response. This procedure was repeated for each Beanie Baby™ (4 for young children, 8 for old children).

Training Phase 2. This phase was the same as Training Phase 1 except that the second set of four or eight Beanie Babies™ was presented to the child using the same procedures.

Testing Phase 2. This phase was the same as Testing Phase 1 except that the second set of Beanie Babies™ was used.

Long-Term Memory Test. About two hours after the completion of Testing Phase 2, the child was tested a second time, in order to assess long-term retention of the Beanie Baby™ names that were presented during training. For the long-term memory test, Testing Phase 1 and Testing Phase 2 were repeated again without any retraining or feedback.

Parent Questionnaire. The names we selected for the Beanie Babies™ were real English words that the children may have known prior to participating in the experiment. It is possible that prior knowledge of the words may play a role in children's ability to associate them with the Beanie Babies™ during the course of the experiment. Names that are entries in the lexicon and are already known may be easier to learn in this procedure because their phonological representations are already stored in memory and so do not require learning a new sound pattern. To evaluate this possibility, we wanted to know which test words each child had prior knowledge of and how familiar the child was with each word. The parents of each child were given a questionnaire about their child's knowledge and familiarity with the words used in the experiment. For each word (e.g., "Red"), the parents were asked to indicate whether "you think your child knows what the word means when he or she hears it spoken (yes or no)". They were also asked to rate, on a scale of 1 to 5, "How familiar your child is with the spoken versions of each word (1: not familiar, 5: very familiar). In other words, to what extent do you feel he or she would recognize each word as something he or she has heard before?" The scores obtained from the parent questionnaires were used to compare children's knowledge and familiarity with the words to their performance with those words on the receptive and expressive tests.

Results

Immediate Tests. The mean proportion of correct responses (out of 8 trials for the young groups and out of 16 trials for the old groups) on the receptive and expressive tests was calculated separately for each child and then averaged for each group. A summary of the results is displayed in the four panels in Figure 1. The left-hand panels display the young children's mean proportion correct response while the right-hand panels display the old children's mean proportion correct response. The upper panels display the scores for the CI children; the lower panels display the scores for the NH children. Within each panel, the mean scores of the receptive and expressive tests are displayed separately. The dark bars in each panel represent the mean proportion correct response on the immediate test; the light bars represent the mean proportion correct response on the delayed test.

The proportion correct response scores were subjected to a four-way repeated-measures ANOVA with Group (CI, NH) and Age (young, old) as between-subjects factors and with Test (receptive, expressive) and Delay (immediate, delay) as within-subjects factors. The ANOVA revealed a significant main effect of Group (F (1, 44) = 132.61, $p < .001$), reflecting higher overall scores for the NH children than the CI children. There was also a significant main effect of Test (F (1, 44) = 82.67, $p < .001$). Overall, the scores were higher on the receptive tests than the expressive tests. CI children demonstrated a larger difference in performance on the receptive versus expressive tests than the NH children. This interaction was significant (F (1, 44) = 7.44, p < .01). Similarly, young children displayed a larger effect of test condition than old children (F (1, 44) = 10.15, p < .01). The Group X Age cross-over interaction approached statistical significance (F (1, 44) = 3.23, p = .08). Young CI children performed better than old CI children, whereas old NH children performed better than young NH children. None of the other main effects or interactions was significant (F values < 1).

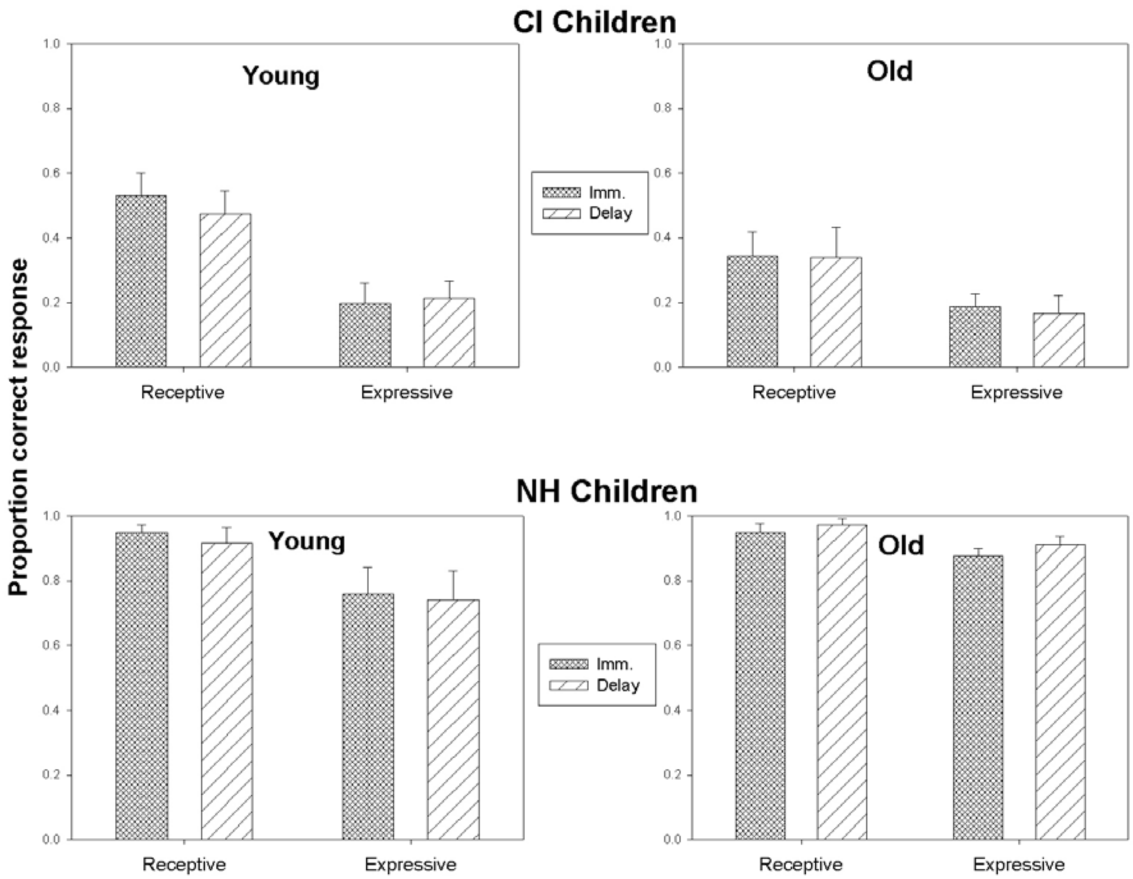


Figure 1. Mean proportion correct for all words for CI and NH children.

Word Familiarity and Word Learning. We selected the words for this experiment so that they would correspond to salient perceptual attributes of the Beanie Babies™ in order to facilitate rapid word learning in the deaf children who use CIs. We expected that these children would only be able to use this information if they had some prior knowledge and familiarity of the words in the first place. To test the role of prior word knowledge on the deaf children's ability to learn the names we correlated their word familiarity scores obtained from the parent questionnaires to their proportion correct responses on the word-learning task. For each word, we calculated an average familiarity score by taking the mean of the familiarity ratings on the parent questionnaires. Then, for each word, we calculated a proportion correct score by taking the mean proportion correct for that word separately for each test condition. Both average familiarity and proportion correct scores were calculated separately for the young and the old CI groups.

Figure 2 displays scatter plots showing the proportion correct for each word in the receptive and expressive tests plotted against the word's average familiarity score. The left-hand panels display the plots for the immediate testing conditions; the right-hand panels display the plots for the delayed testing condition. Scatter plots for the young CI group are shown in the upper panels while those for the old CI group are shown in the lower panels. Correlations are summarized in Table 4.

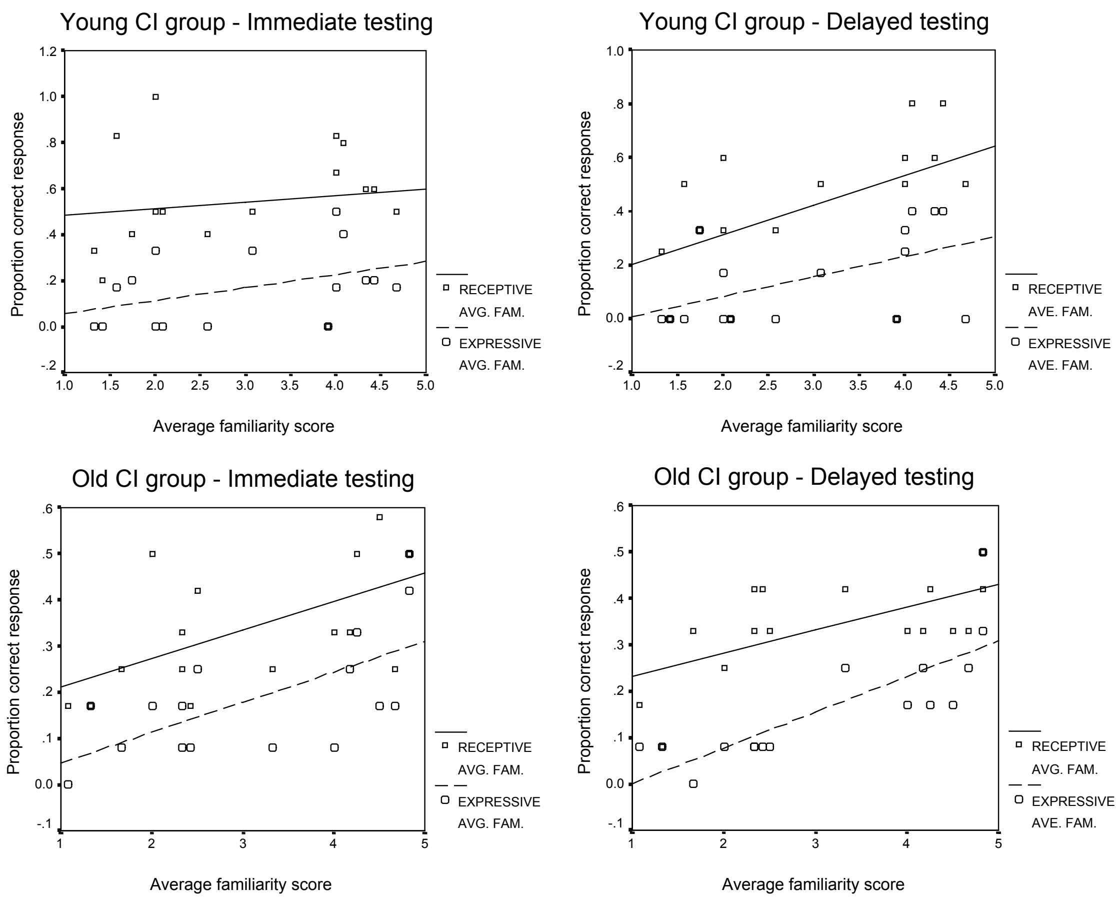


Figure 2. Scatter plots of word analyses for CI children, separated by age group and testing condition.

Score	1	2	3	4	5
CI Young Group					
1. Average familiarity rating	–	.13	.53*	.43	.53*
2. Receptive immediate mean score		–	.76**	.46	.33
3. Receptive delay mean score			–	.62*	.66**
4. Expressive immediate mean score				–	.71**
5. Expressive delay mean score					–
CI Old Group					
1. Average familiarity rating	–	.60*	.63**	.64**	.81**
2. Receptive immediate mean score		–	.41	.66**	.43
3. Receptive delay mean score			–	.44	.47
4. Expressive immediate mean score				–	.69**
5. Expressive delay mean score					–

Table 4. Item correlations between average familiarity ratings of stimulus name and proportion correct response for receptive and expressive tests.

For the Young CI group, average familiarity was positively correlated with both the receptive (r = +.53, $p < .05$) and expressive tests (r = +.53, $p < .05$) in the delayed condition. Although the correlations were in the same direction for the receptive (r = +.13, p = .62) and expressive (r = +.43, p = .10) tests in the immediate conditions, they were not statistically significant. For the old CI group, average familiarity was correlated significantly with both tests in the immediate testing condition (receptive: r = +.60, $p < .05$; expressive: r = +.64, $p < .01$) and for both tests in the delayed testing condition (receptive: r = +.63, $p < .01$; expressive: r = +.81, $p < .01$). These correlations suggest that prior word knowledge produced effects on word learning in this task.

Analyses with Word Knowledge as a Factor. The item correlations of word familiarity with word-learning test scores showed a relation between CI children's prior familiarity with the stimulus words and their performance on the word-learning tests, especially in the delay conditions. To analyze these findings further, we recalculated the CI children's scores, separating their performance on words they knew prior to testing from words they did not know before they began the study, based on information in the parent report questionnaire.[7]

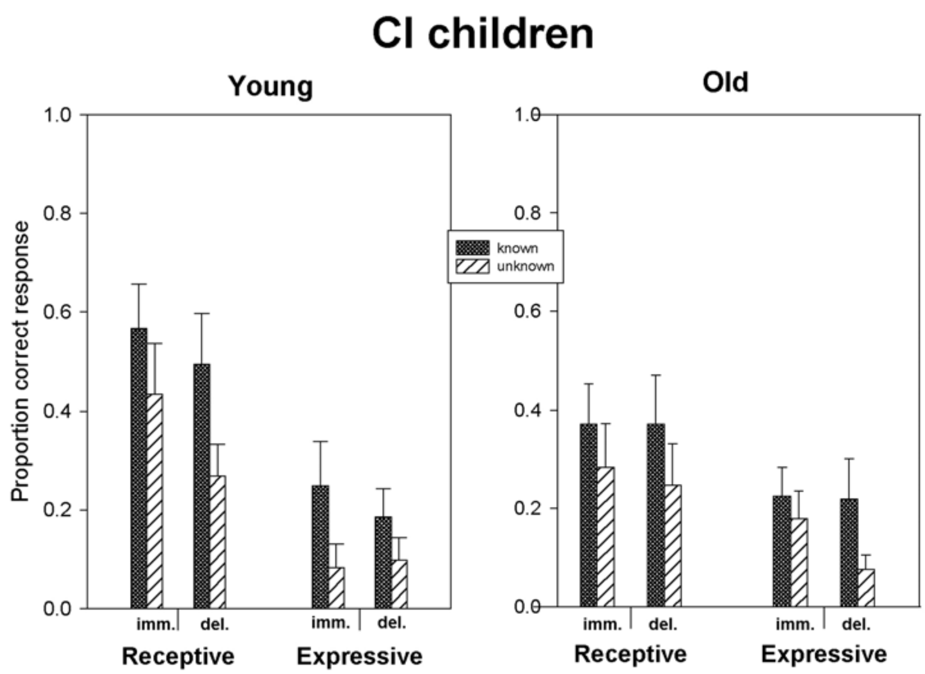


Figure 3. Mean proportion correct for CI children separated by whether they had prior knowledge of the words.

Figure 3 displays the young CI children's performance in the left-hand panel and the old CI children's performance in the right-hand panel. As in Figure 1, the scores within each panel are separated by test type and delay condition. However, in this figure the data are further separated by scores for words that were known (dark bars) versus scores for words that were unknown (light bars). The proportion correct responses for known and unknown words were then subjected to a four-way repeated-measures

[7] We separated the scores based on word knowledge rather than word familiarity because familiarity would have been a factor with five levels, and some of the levels would have been empty for many of the children. Word knowledge scores were highly correlated with word familiarity scores (r = .97, $p < .001$).

ANOVA with Age as the between-subjects factor and Test, Delay, and Word Knowledge (known, unknown) as within-subjects factors. As in the first analysis, performance on the receptive tests was higher than on the expressive tests (F (1,22) = 56.00, p < .001). We also observed a significant Test (receptive vs. expressive) X Age (young vs. old) interaction (F (1,22) = 6.21, p < .05), reflecting a larger difference in performance across tests for the younger children than for the older children. Unlike the first analysis, however, we found a significant main effect of Delay (F (1,22) = 5.41, p < .05), reflecting higher scores on the immediate than the delayed testing condition. Finally, the CI children performed significantly better on known than on unknown words (F (1,22) = 12.54, p < .01). No other main effects or interactions were statistically significant.

Re-analysis with Known Words Only. In the previous analysis, we found a significant main effect of word knowledge on learning with the CI children. Higher scores were observed for known words than unknown words. According to the parent questionnaires, the NH children knew an average of 95% of the words whereas the CI children knew only 60% of the words. It is possible that the difference in performance observed earlier between the CI and NH children was based on prior word knowledge rather than differences in word learning and encoding skills acquired during the course of the present experiment. To assess this possibility, we reanalyzed the proportion correct scores of the CI and NH children using only the scores for the known words.

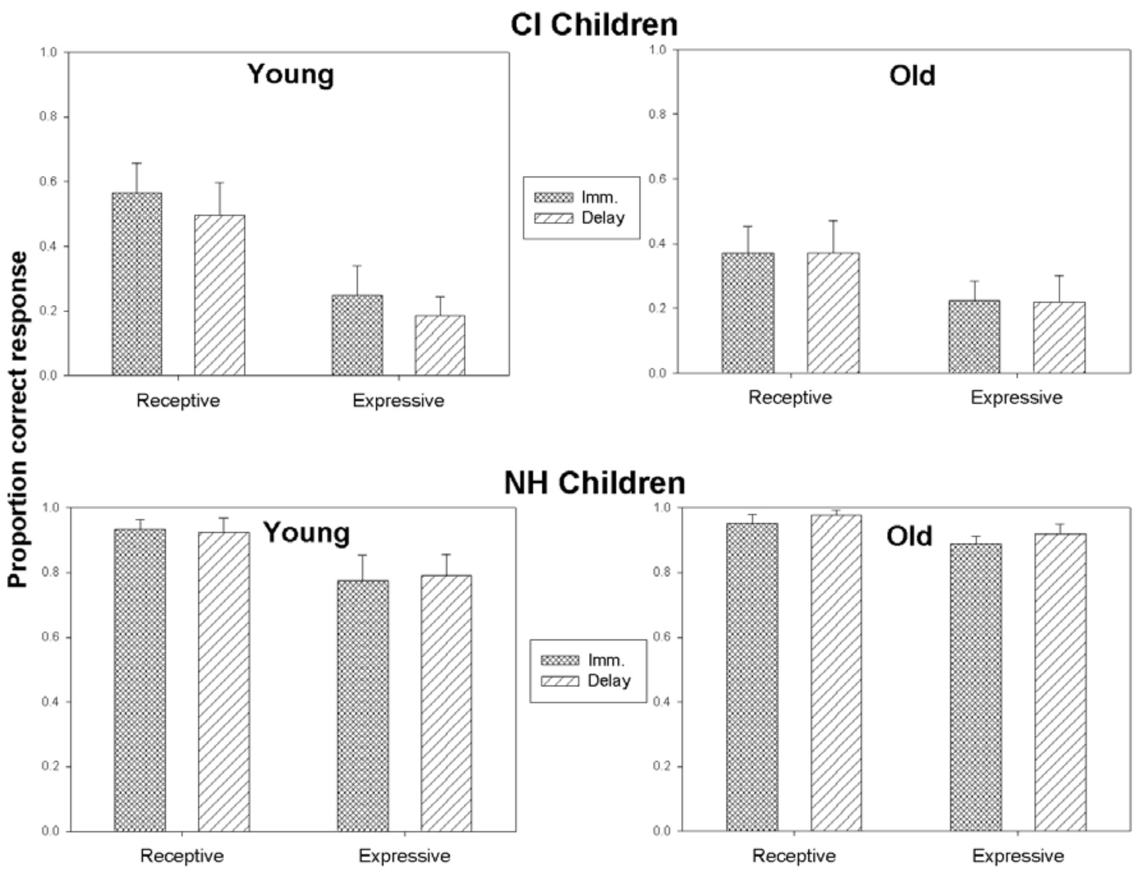


Figure 4. Mean proportion correct of known words only for CI and NH children.

Figure 4 displays the same test conditions as Figure 1 but now only data from the known words is shown. The scores for the known words were subjected to a four-way repeated-measures ANOVA with Group and Age as between-subjects factors and with Test and Delay as within-subjects factors. Overall, the pattern of results was quite similar to the original analyses. In the first analysis, the NH children out-performed the CI children (F (1,44) = 92.17, p < .001) and both groups of children performed significantly better on the receptive than the expressive tests (F (1,44) = 68.66, p < .001). Also, inspection of the interactions revealed that the differences between the test conditions (receptive vs. expressive) were significantly greater for the CI children than the NH children (F (1,44) = 10.129, p < .01) and greater for young children than old children (F (1,44) = 9.54, p < .01). As in the earlier analysis, the Group X Age cross-over interaction for known words did not approach statistical significance, although it was in the same direction (F (1,44) = 1.80, p = .19), reflecting better performance by the young CI and old NH groups. All other main effects and interactions were not significant (F values < 2).

Correlations with Demographics. In order to assess the relationship of performance on the word-learning tests with demographic measures, we correlated the mean proportion correct responses with chronological age, age at implantation, and length of CI use. A summary of these correlations is shown in Table 5. As can be seen in this table, significant negative correlations were observed between age at implantation and scores on the immediate receptive tests, suggesting that children who received cochlear implants at younger ages, and hence had a shorter periods of auditory deprivation, tended to perform better on this receptive test of word learning than children who received their implants at older ages. The correlations between age at implantation and the other tests were in the same direction but were not statistically significant. None of the other correlations were significant.

	1	2	3	4	5	6	7
1. Chronological age	–	.86**	.46*	-.30	-.30	-.02	-.13
2. Age at implantation		–	-.07	-.42*	-.32	-.17	-.22
3. Length of CI use			–	.15	-.03	.26	.14
4. Receptive immediate mean score				–	.75**	.57**	.76**
5. Receptive delay mean score					–	.61**	.79**
6. Expressive immediate mean score						–	.64**
7. Expressive delay mean score							–

Table 5. Correlations between mean proportion correct responses and chronological age, age of implantation, and length of CI use for all children with CIs.

General Discussion

The aim of this study was to develop a procedure that could be used to investigate the word-learning skills of profoundly deaf children who have received CIs. We are interested in understanding the cognitive and linguistic factors that contribute to differences in audiological outcome and language development that have been found between deaf children who use CIs and NH children and the high degree of variability in the linguistic skills of profoundly deaf children following cochlear implantation. A word-learning task using Beanie Babies™ was developed and implemented to study the receptive and expressive skills of deaf children. This study represents the first investigation of novel “word” learning skills in this clinical population and provides an important initial step in understanding the fundamental mechanisms and building blocks of language development after a period of auditory deprivation due to deafness and the restoration of hearing following cochlear implantation.

All of the children who were tested in the procedure were able to complete the task. Only four of the deaf children and none of the NH children performed at chance levels on all of the testing conditions. Thus, the training and testing procedures we developed appeared to be at a level of difficulty appropriate for testing these two populations of children. We also observed a wide range of performance across children, especially among the deaf children. For the immediate receptive test, the range of scores varied from 6% to 100% correct for the deaf children and from 69% to 100% for the NH children. Similar variation was found for the delayed receptive test. For the expressive tests, the ranges for the deaf children were from 0% to 63% for both immediate and delay conditions. We also found a wide range of scores on the expressive tests for the NH children as well (25%-100% immediate; 0%-100% delay). However, the lower ends of the ranges were due to two of the NH children who did not give many responses at all, presumably because of shyness. All other NH children performed at or above 50%. The median scores for the immediate and delay expressive tests were 88% and 92.5%, respectively.

In addition to individual variability on the word-learning measures among deaf children who use CIs, we also found significant differences in performance between the two groups of children. NH children performed significantly better than the deaf children on both receptive and expressive tests of word learning at both immediate test and under delayed conditions. These findings are not surprising and are consistent with the results of other investigations which have found that while CIs improve spoken language skills of deaf children, these implanted children generally do not display linguistic skills that are comparable to NH children (Svirsky et al., 2000). The present findings suggest that hearing impairment not only affects speech perception but it also affects cognitive and linguistic skills that are important for novel word learning. There are several possible ways in which early auditory deprivation and hearing impairment can affect the development of word-learning skills, including impairments in phonological processing and phonological working memory. These possibilities will be considered later.

In addition to finding differences in word-learning performance between deaf and hearing children, we also observed a main effect for the type of test. Both NH children and deaf children performed significantly better on the receptive word-learning test than on the expressive test. The pattern of results was the same for both groups under the delay conditions. The difference in performance between receptive and expressive tasks is not surprising because there are fundamental differences in the information processing demands of these two kinds of tasks. In both procedures, performance depends on the child's ability to rapidly encode the spoken names into memory and then learn to associate them to the correct Beanie Baby™ referent. Expressive tests are generally more difficult because they require the child to form mental representations of the names that are robust enough to access from memory without any explicit retrieval cues or context. In receptive tests, on the other hand, phonological and lexical representations of the correct names can be accessed more easily from memory because the experimenter provides retrieval cues and context at the time of the test. The child only has to recall the association between the given name and the corresponding Beanie Baby™ and does not have to construct a phonological representation "on-the-fly" without cues or context.

The deaf children also showed larger differences in performance between receptive and expressive tests than the NH children. It is possible that some of the representations of the names that deaf children formed were not robust or well specified and were thus difficult to access from the mental lexicon although they could be recognized in the receptive tests. If this account is correct, then we might expect that words that were identified successfully on the expressive tests would also tend to be recognized on the receptive tests. To assess this possibility, we calculated the proportion of correctly identified words in the expressive tests that were also correctly recognized in the receptive tests. For the younger children, 91% of items that were responded to correctly during the expressive tests were also responded to correctly during the receptive tests. For the older children, the proportion was 76%. These results are consistent with the proposal that the deaf children were able to form robust representations for some of the names and these names were responded to correctly in both the receptive and expressive tests.

The extent to which deaf children are able to form robust representations that are easily retrievable from memory may be crucial for their ability to learn novel words. To assess the robustness of the children's lexical representations, we also measured retention of the names several hours after the immediate tests. The initial ANOVA on all of the stimuli using both known and unknown words combined revealed no effect of delay, suggesting that the word representations that were formed during the acquisition phase were robust enough to persist in long-term memory. However, a re-analysis based on only the words that children knew prior to the experiment also failed to reveal any effects of delay. Other analyses demonstrated that 69% of responses that were correct in immediate testing were also correct in delay testing for the younger group. For the older group of children, the proportion was 73%. These findings suggest that once a name was successfully learned and encoded in memory during the training phase, the particular name was retained even after a delay of several hours.

Our initial analyses were calculated using either all of the test words or only a subset of the words that the children had previous experience with before the experiment. In both of these cases, we found no effects of delay on performance. However, two interesting patterns emerged when we examined the performance of the deaf children on words they knew before the experiment compared to unfamiliar words that were previously unknown to them. First, we found that the CI children performed significantly better on known than unknown words. Second, we found that the deaf children performed significantly worse on the delayed tests than on the immediate tests with the words that were previously unknown. In contrast, they did not perform significantly worse across delay conditions for the words that were previously known. These results suggest that deaf children had more difficulty encoding novel sound patterns than familiar sound patterns in these name-learning tasks. NH children did not display the same pattern of results.

Given the absence of an effect of delay for both groups of children on familiar words, it is possible that the deaf children may not have demonstrated any novel word learning at all. Instead, they may have simply selected the animals based on their prior knowledge of the word attributes without consideration as to how the objects were named during the training. While the deaf children may have used their prior knowledge of the attribute names to help them in narrowing down their response choices, there are at least two reasons why we think that learning during the training phase played an important role in the children's responses. First, the deaf children did not perform near ceiling for any of the familiar words and their performance on unfamiliar words was not at floor. Thus, word knowledge per se was not sufficient for generating a correct response, and not knowing a word did not always result in failure to learn the attribute names of the Beanie Baby™ stuffed animals. Second, if the deaf children relied solely on their prior knowledge of attribute names to select the Beanie Babies™ during the tests, we would expect them to perform much better with the animals that had unique attribute names than the animals whose attribute names could have been applied to at least one other Beanie Baby™ in the test set. However, this did not occur. The mean familiarity score for the deaf children on the uniquely named words was 3.09 and the mean proportion correct across the tests for those words was 0.29. Similarly, the mean familiarity score for the deaf children on the confusable names was 3.03 and the mean proportion correct for those was also 0.29. Clearly, the uniqueness of the attribute names did not appear to play a role in these children's ability to select the correct animals from a set or to recall the names of the animals. Thus, while prior word knowledge and familiarity did facilitate CI children's ability to match the names to the correct Beanie Babies™, they apparently did not rely solely on their prior knowledge of the attribute names to select the Beanie Babies™ during the tests. In other words, novel word learning took place during the training phase of the experiment.

Several other factors may have also contributed to the variability in word-learning performance among the deaf children and to the overall differences in performance between the two groups of children. It is well known that hearing loss and early auditory deprivation lead to impairment in the ability to process and encode phonological information. Deaf children may have had more difficulty in the word-learning tasks than the NH children simply because they were not able to encode and store the phonological information and form representations of the sound patterns of the words. However, this possibility does not seem likely because all of the children were able to repeat the words back to the experimenter, in some form or another immediately after presentation, during both the study and acquisition phases of the procedure. Also, we observed a larger difference between receptive and expressive tests for the deaf children than the NH children. Performance of the deaf children on the receptive tests was much better than on the expressive tests. These results suggest that the deaf children were able to form coarse phonological representations of the sound patterns of the names, but these representations may not be detailed and robust enough to access them and use them in an expressive recall task. Furthermore, the deaf children demonstrated an effect of delay with the unknown words, which suggests that they formed lexical representations, but that these representations were not robust enough to be retained in long-term memory over a period of a few hours.

Another factor that may be responsible for differences in performance between the two groups of children is working memory capacity. According to Baddeley's (1986) model, one of the components of working memory is the phonological loop that maintains phonological representations for a short period of time. The phonological representations of words enter the phonological loop through the "inner ear" or the "inner voice" and are maintained there by repeating the words to the inner ear either by repetition from an external source or by the inner voice rehearsing the words. Investigations of working memory in NH children have found that children who have larger working memory capacities also tend to have better word-learning skills and larger vocabularies (e.g., Gathercole & Baddeley, 1989; Gupta & MacWhinney, 1997). Baddeley and his colleagues have suggested that working memory serves as an important mechanism for word learning – the phonological loop is a temporary store for novel sound pattern of words until the words can be encoded and stored into long-term memory (Baddeley, Gathercole, & Papagno, 1998).

In a recent paper, Pisoni and Cleary (in press) summarized a series of recent experiments that investigated the working memory and word recognition skills of deaf children following cochlear implantation. They measured 8- and 9-year-old deaf CI users' immediate memory spans using the WISC-III digit span task and assessed their word recognition skills using both open and closed-set word recognition tests. Pisoni and Cleary found strong positive correlations between performance on the forward digit span task, which is assumed to measure the capacity of verbal short-term memory and performance on several word recognition tasks. They also tested a group of age-matched NH children on the forward digit span task and found that these children performed better on the forward digit span task than the deaf children with CIs. Pisoni and Cleary's findings suggest that the word recognition skills of deaf children who use CIs are closely associated with the phonological coding of the sound patterns of words and working memory, which appears to be "atypical" and somewhat delayed in deaf children after cochlear implantation compared to age-matched NH children.

The present investigation did not directly assess the children's working memory capacity or immediate memory span. However, several of our findings are consistent with the hypothesis that the deaf children have atypical verbal working memory capacity relative to the NH children. For example, the poorer performance after a period of delay with unknown words suggests that the deaf children had more difficulty encoding novel unfamiliar words into long-term memory, which suggests impairment in the operation of their phonological loop and verbal rehearsal mechanisms (Baddeley et al., 1998). One reason why deaf children may have more difficulty forming robust lexical representations of novel words than

NH children is that they may not be able to verbally rehearse and maintain phonological patterns as efficiently in working memory so that they can be encoded into long-term memory (see Cleary, Dillon, & Pisoni, 2002).

There are several reasons why phonological processing and/or working memory may be impaired in deaf children following cochlear implantation. As mentioned earlier, deaf children may have difficulty encoding phonological information because their auditory processing skills are poorer relative to NH children in terms of frequency discrimination and temporal coding of the speech waveform. However, there may be factors other than their auditory capacities and discrimination skills that may play a role in processing, encoding, and storing phonological information. It is important to remember that prior to cochlear implantation these children had a profound hearing loss and experienced a period of sensory deprivation and lack of auditory stimulation. Neural and cognitive development developed and unfolded in the absence of auditory stimulation during critical periods in development (Shepherd, Hartmann, Heid, Hardie, & Klinke, 1997). One consequence of early auditory deprivation appears to be a diminished ability to selectively attend to auditory information due to lack of stimulation during early stages of development. Attention to sound and especially speech sounds is critical for processing and encoding phonological information and it is even more important for the development of working memory and verbal rehearsal processes used in word learning and lexical development. If attention to speech sounds is not highly automatic, then words may not be as readily encoded into the phonological loop for rehearsal and eventual storage into long-term memory.

Another factor that may contribute to the development of phonological skills that are important for word learning is vocabulary size. As we noted in the introduction, Lederberg et al. (2000) found that deaf and hard-of-hearing children with larger vocabularies displayed more advanced word-learning skills than deaf and hard-of-hearing children with smaller vocabularies. They speculated that some word-learning skills, such as implicit fast mapping skills, might emerge only when vocabulary size becomes large enough. It is possible that deaf children who have received CIs may not develop typical phonological processing and encoding skills necessary for the implementation of fast mapping strategies until they have acquired a sufficiently large vocabulary of words in their mental lexicons.

The present study was designed to investigate the novel word-learning skills of deaf children following cochlear implantation. We found that deaf children who use CIs have more difficulty learning names after only a few exposures than NH children but that they do not necessarily exhibit an impaired ability to retain familiar words in long-term memory. One of the issues raised by these findings is how vocabulary size is related to deaf children's ability to learn novel words. Because large vocabularies reflect more advanced word-learning skills (Lederberg et al., 2000), it would be worth exploring further the underlying lexical processes that contribute to individual differences of vocabulary size in this clinical population. One possibility is that there may be large differences in the quality of the acoustic-phonetic input that CIs deliver for each child, resulting in differences in auditory acuity. Some deaf children may be able to form quite detailed phonological and lexical representations from the auditory input provided by their CIs while other deaf children may only be able to form coarser representations that are easily confused with other phonetically similar words. One way to address this possibility would be to conduct a novel word-learning experiment using sets of nonwords that are specifically designed to have different degrees of phonological confusability.

Another factor that may affect vocabulary size is the deaf children's phonological processing skills and working memory capacity. Even if deaf children are able to discriminate fine-grained acoustic-phonetic properties in speech, they may be unable to build a vocabulary and develop a lexicon of words if they are not able to initially encode these sound patterns into working memory and store them in long-term memory. Gilbertson and Kamhi (1995) found that for children with mild to moderate hearing-losses, measures of their phonological processing skills predicted their subsequent word-learning abilities. Similarly, Gathercole and Baddeley have found that working memory is an important factor for building a vocabulary (Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1989; 1990). Obtaining several measures of a deaf child's phonological processing skills (i.e., phoneme identification, phoneme monitoring, nonword repetition, etc.) and working memory capacities would help in understanding the relations between phonological coding, working memory and novel word learning in this clinical population.

In the present experiment, deaf children, like the NH children, did not exhibit any decline in performance on the receptive and expressive word-learning tasks across delay conditions when the words they had to learn were familiar. However, for unfamiliar words, deaf children performed worse on the word-learning tasks after a delay than during the immediate testing. These findings suggest that the long-term memory of deaf children who use CIs may be impaired relative to that of NH children. Markson and Bloom (1997) reported that NH children were able to recall novel nonwords after a 1-month delay, suggesting that NH children's long-term memory for the sound patterns of novel words is very robust. Similar novel word-learning studies with longer delay periods would be useful for investigating the long-term encoding and memory capacities of deaf children who use CIs.

Several traditional demographic factors may also contribute to individual differences in performance on word-learning tasks among deaf children who use CIs. These factors include age at implantation, length of CI use, communication mode, length of deafness before implantation and number of active electrodes. In the present study, we found weak correlations between age at implantation and performance on word-learning tasks. However, we intentionally selected children in this study who came from oral-only educational environments and had at least 1 year of CI experience so that they would be able to complete these auditory tasks. A better assessment of the contribution of the demographic factors will require investigations with groups of children in which there is more variance in these demographic factors across children.

Finally, it is important to investigate the effects of early auditory deprivation on novel word-learning skills of deaf children after intervention with a cochlear implant. Some researchers have noted that profoundly hearing-impaired children may develop phonological processing deficits that are similar to children who have been diagnosed with Specific Language Impairment. Indeed, some investigators have considered the possibility that hearing impairment may be a risk factor for SLI (Gilbertson & Kamhi, 1995), and one of the factors observed in SLI is difficulty with phonological processing skills (Leonard, 1998). Thus, it is possible that both hearing-impaired children and children with SLI have similar difficulties in encoding speech sounds and maintaining the phonological pattern of words in both short- and long-term memory. In particular, both groups of children may have difficulty with rapid learning of words.

Support for this hypothesis comes from a series of studies by Rice and her colleagues who report that children with SLI have difficulty with "quick and incidental learning" which is similar to fast mapping (Oetting, Rice, & Swank, 1995; Rice, Buhr, & Nemeth, 1990; Rice, Oetting, Marquis, Bode, & Pae, 1994). In one investigation, Rice et al. (1994) found that children with SLI needed many more exposures to words than normally developing children in order to display even a basic understanding of the words. Moreover, children with SLI appeared to be more likely to forget the meanings of words after a short delay, suggesting that their long-term memory retention of words and word meanings was atypical as well. In sum, deaf children who use CIs may have difficulties quickly learning novel words due to poorly developed phonological processing skills, and these differences may be similar to the results found with children who have SLI. Further novel word-learning studies with deaf children who use CIs and hearing-impaired children who use hearing aids may provide further insights into how hearing impairment affects language development and, especially how degraded sensory information influences novel word learning and vocabulary development.

Conclusions

We found that deaf children who received CIs are able to perform a novel word-learning task in which they were exposed to names for Beanie Babies™ and then tested for their receptive and expressive knowledge of these names. We observed a high degree of variability in performance in this task, suggesting that some of the underlying phonological skills needed for novel word learning are delayed or disordered in this population of children.

Deaf children who use CIs performed worse than NH children in all test conditions, but neither group exhibited a decline in performance across delay conditions for familiar words that they knew before testing began. Only when we examined the learning of unfamiliar words did we see that deaf children's performance declined across delay. These findings suggest that the poorer performance on novel word-learning tasks of deaf children with CIs may be due to differences in phonological processing or verbal working memory capacity than storage or retrieval from long-term memory. However, additional research is needed to further explore the relations between in phonological processing skills, working memory capacity, and long-term retention in deaf children who use CIs and their contributions to novel word learning.

References

Baddeley, AD. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Briscoe, J., Bishop, D.V.M., & Norbury, C.F. (2001). Phonological processing, language, 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.

Carey, S. (1978). The child as word learner. In M. Halle & J. Bresnan & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264-293). Cambridge, MA: MIT Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development 15*. (pp. 17-29). Stanford University.

Clark, E.V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 65-110). New York: Academic Press.

Clark, E.V. (1983). Meanings and concepts. In J.H. Flavell & E.M. Markman (Eds.), *Cognitive Development* (Vol. III, pp. 787-840). New York: Wiley.

Cleary, M., Dillon, C., & Pisoni, D.B. (2002). Imitation of nonwords by deaf children after cochlear implantation: Preliminary findings. *Annals of Otology, Rhinology, & Laryngology, 111 (Suppl. 189),* 91-96.

Eisenberg, L.S., Martinez, A.S., Sennaroglu, G., & Osberger, M.J. (2000). Establishing new criteria in selecting children for a cochlear implant: Performance of "platinum" hearing aid users. *Annals of Otology, Rhinology, and Laryngology, 109 [Part 2 Suppl.]*, 30-33.

Fenson, L., Dale, P., Reznick, S., Bates, E., Thal, D., & Pethick, S. (1994). *Variability in Early Communicative Development. Monographs of the Society for Research in Child Development, 59* (Serial number 242).

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

Gathercole, S.E., & Baddeley, A.D. (1989). Development of vocabulary in children and short-term phonological memory. *Journal of Memory and Language, 28*, 200-213.

Gilbertson, M., & Kamhi, A.G. (1995). Novel word learning in children with hearing impairment. *Journal of Speech and Hearing Research, 38*, 630-642.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore: Paul H. Brookes Publishing Co.

Heibeck, T.H., & Markman, E.M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*, 1021-1034.

Houston, D.M., & Jusczyk, P.W. (submitted). *Infants' long-term memory for words and voices*. Manuscript submitted to *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236-248.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1-23.

Jusczyk, P.W., & Hohne, E.A. (1997). Infants' memory for spoken words. *Science, 277,* 1984-1986.

Kirk, K.I. (2000). Challenges in the clinical investigation of cochlear implant outcomes. In J.K. Niparko (Ed.), *Cochlear implants: Principles & practices* (pp. 225-259). Philadelphia: Lippincott, Williams & Wilkins.

Kirk, K.I., Osberger, M.J., Robbins, A.M., Riley, A.I., & Todd, S.L. (1995). Performance of children with cochlear implants, tactile aids, and hearing aids. *Seminars in Hearing, 16*, 370-381.

Lederberg, A.R., Prezbindowski, A.K., & Spencer, P.E. (2000). Word-learning skills of deaf preschoolers: The development of novel mapping and rapid word-learning strategies. *Child Development, 71,* 1571-1585.

Lederberg, A.R., & Spencer, P.E. (2001). Vocabulary development of deaf and hard of hearing children. In M.D. Clark & M. Marschark (Eds.), *Context, cognition, and deafness* (pp. 88-112). Washington, DC: Gallaudet University Press.

Leonard, L. B. (1998). *Children with Specific Language Impairment*. Cambridge, MA: MIT Press.

Loizou, P.C., Dorman, M., & Tu, Z.M. (1999). On the number of channels needed to understand speech. *Journal of the Acoustical Society of America, 106*, 2097-2103.

Markman, E.M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on Language and Thought* (pp. 72-106). Cambridge: Cambridge University Press.

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature, 6619*, 813-815.

Nelson, K. (1988). Constraints on word learning? *Cognitive Development, 3*, 221-246.

Oetting, J.B., Rice, M.L., & Swank, L.K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech and Hearing Research, 38*, 434-445.

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

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.), *Handbook of auditory research: Auditory prostheses* (Vol. 10). New York: Springer-Verlag.

Pisoni, D.B., Cleary, M., Geers, A.E., & Tobey, E.A. (2000). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *Volta Review, 101*, 111-164.

Pisoni, D.B., Svirsky, M., Kirk, K. I., & Miyamoto, R.T. (1997). Looking at the "Stars": A first report on the intercorrelations among measures of speech perception, intelligibility and language development in pediatric cochlear implant users. *Progress Report on Spoken Language Processing #21* (pp. 51-91). Bloomington, IN: Indiana University, Department of Psychology.

Rice, M. L., Buhr, J., & Nemeth, M. (1990). Fast mapping word learning abilities of language delayed preschoolers. *Journal of Speech and Hearing Disorders, 55*, 33-42.

Rice, M.L., Oetting, J.B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on SLI children's word comprehension. *Journal of Speech and Hearing Research, 37*, 106-122.

Shepherd, R.K., Hartmann, R., Heid, S., Hardie, N., & Klinke, R. (1997). The central auditory system and auditory deprivation: Experience with cochlear implants in the congenitally deaf. *Acta Oto-laryngologica, Suppl 532,* 28-33.

Svirsky, M., Robbins, A., Kirk, K.I., Pisoni, D.B., & Miyamoto, R.T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science, 11*, 153-158

Tobey, E., Geers, A.E., & Brenner, C. (1994). Speech production results: Speech feature acquisition. *Volta Review, 96*, 109-130.

Tomblin, J.B., Spencer, L., Flock, S., Tyler, R.S., & Gantz, B.J. (1999). A comparison of language achievement in children with cochlear implants and children using hearing aids. *Journal of Speech, Language, and Hearing Research, 42*, 497-511.

Zwolan, T.A., Zimmerman-Phillips, S., Ashbaugh, C.J., Hieber, S.J., Kileny, P.R., & Telian, S.A. (1997). Cochlear implantation of children with minimal open-set speech recognition skills. *Ear & Hearing, 18*, 240-251.

Appendix

Sample Scenario:

This is *Name.*
Can you say hi to him?
Say "Hi *Name!*"
Now your turn {child says "hi *Name*"}
Name likes to climb the tree.
Can you put him on the tree? {child interacts with BB}
Look – *Name* is on the tree.
Tell him to get down. {child says, "get down *Name*"}
Good. Now, *Name* has to go bye bye.
Say, bye bye *Name.* {child repeats "bye bye *Name*"}

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

Using Immediate Memory Span to Measure Implicit Learning[1]

Jeffrey Karpicke

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

[1] This research was supported by NIH Research Grant DC00111 and Training Grant DC00012 to Indiana University. I would like to thank Miranda Cleary and Winston Goh for their help with several of the technical aspects of this project. I must especially thank Luis Hernandez, who contributed to several pivotal discussions, and Lorin Lachs, who has been an additional mentor to me throughout the past two years. I want to acknowledge my honors thesis committee members Dale Sengelaub, Tom Busey, and my thesis advisor, David Pisoni, whom I especially thank for having invested such a great deal of personal interest in my intellectual development over the past two years.

Using Immediate Memory Span to Measure Implicit Learning

Abstract. Implicit learning is the process of acquiring knowledge about the structure of a complex stimulus environment largely independent of conscious awareness of either the process or the products of acquisition. In the present study, a new method of measuring implicit learning using immediate memory span capacity was developed. This method avoids many of the conceptual and methodological pitfalls found in standard implicit learning tasks. Subjects were presented with sequences that had been generated by an artificial grammar and were asked to reproduce the sequences by pressing buttons on a customized response box. After being exposed to these sequences, subjects showed a selective increase in memory span for novel sequences governed by the same grammatical constraints, despite being largely unable to display explicit knowledge about grammatical sequences in a recognition memory test. Individual differences in the degree of implicit learning co-varied with general measures of memory span, indicating that individuals with larger immediate memory capacity are better able to learn and subsequently exploit the higher-order sequential dependencies that compose the structure of an artificial grammar. Implications for current models of implicit learning, the relationships between immediate memory span and sequential learning, and individual differences in implicit learning and language processing are discussed.

Introduction

Implicit learning has been defined as the process by which knowledge about the structure of a complex stimulus environment is acquired largely independent of conscious awareness of either the process or the products of acquisition (Reber, 1997a). This mechanism of automatic, non-conscious, non-reflective, non-analytic absorption of information from the environment is ubiquitous in human experience. Implicit learning has been thought of as a general, domain-free learning mechanism for inducing regularities about patterned relationships in a stimulus environment. Furthermore, the implicit learning process works in parallel with, not in the absence of, explicit cognitive processes: While explicit attention is paid to highly salient elements in the environment, information about structural regularities underlying those elements can be induced implicitly (Winter & Reber, 1994). Real world examples of implicit learning include learning appropriate social behaviors, complex procedures, and, perhaps most importantly, natural languages. Indeed, interest in the basic mechanisms of language learning led directly to the development of early laboratory methods for studying implicit learning (Chomsky & Miller, 1958; Miller, 1958; Reber, 1967).

This introduction is organized into several sections. The first discusses distinctions and commonalities among implicit and explicit learning and memory, which are related constructs distinguished in terms of the processes engaged in during encoding and retrieval. The development of finite-state languages (i.e., artificial grammars) and the emergence of implicit learning research are discussed. The second section highlights significant studies from the implicit learning literature that surround three central controversies: whether learning in an implicit learning task truly occurs largely independent of conscious awareness, whether the resulting knowledge base is an abstract system of rules or an instance-based set of exemplars, and ultimately whether distinct, dissociable modes of human learning even exist. The third section outlines a model of the cognitive unconscious based on principles from evolutionary biology, proposed by Reber (1992a, 1992b, 1993; also Reber & Allen, 2000), which predicts that implicit cognitive processes are not as susceptible to individual differences as explicit cognitive processes. The few studies that specifically investigate individual differences in implicit learning are then reviewed.

The final section argues that failures to find systematic variations in performance among individuals in implicit learning tasks can be attributed to several conceptual and methodological problems. With the exception of a few recent studies (e.g., Destrebecqz & Cleeremans, 2001), research on implicit learning has widely assumed that tasks are "process-pure" – that a given task exclusively involves either implicit or explicit knowledge. A new method for measuring implicit learning is proposed, based largely on assumptions from the process dissociation procedure (PDP; Jacoby, 1991), for separating the contributions of implicit and explicit processes in a given task. Finally, it is suggested that a measure of individual differences in implicit learning might be useful in explaining and predicting individual differences in cognitive abilities that are assumed to rely on implicit mechanisms (e.g., individual differences in language abilities).

Implicit and Explicit Learning and Memory

Implicit memory can be defined as the facilitation of task performance through prior experiences in the absence of conscious or intentional recollection of those experiences (Graf & Schacter, 1985). More specifically, a test of implicit memory does not make explicit reference back to the original study phase (Schacter, 1987), whereas a test of explicit memory does. A classic example of a test of explicit memory is a traditional recognition memory task, where subjects first study a list of items (most often, a list of words) and later are asked to decide whether a given item occurred on the original study list (e.g., Mandler, 1980). In principle, a recognition memory task requires the subject to consciously recollect some part of the original study phase. This methodology may be contrasted with some classic examples of implicit memory tests, such as word fragment completion and anagram solving. In a word fragment completion task, subjects study a list of words (as they would in a recognition memory task) and later are asked to identify words presented in a degraded or altered form. For example, the word METAL could be presented in "visual noise", with some proportion of the orthography masked or deleted (Warrington & Weiskrantz, 1970); the word could be presented as _e_al, with some of the letters missing (Roediger, Weldon, Stadler, & Riegler, 1992); or the word could be presented as an anagram, such as EMTLA (Jacoby & Dallas, 1981). The general finding from these studies is that prior presentation of a word during the study phase improves subjects' ability to identify the degraded or altered word during the testing phase, a phenomenon known as priming (Tulving & Schacter, 1990). The task of identifying a word fragment or unscrambling an anagram does not require the subject to consciously recollect the original study phase, yet their prior exposure to the target word during the study phase has an implicit influence on their ability to identify the word during the testing phase.

Interest in the dissociations between implicit and explicit memory emerged primarily from the study of the amnesic syndrome, in which any of a number of forms of brain damage can impair memory functioning while leaving other cognitive functions intact (Roediger, 1990b). Landmark research by Warrington and Weiskrantz (1968, 1970) challenged the view that amnesic patients lacked the ability to transfer verbal information from short-term to long-term memory by demonstrating intact long-term memory in amnesic patients on implicit memory tests. Warrington and Weiskrantz (1970) compared amnesic patients and control subjects in their performance on free recall, recognition, word fragment identification, and word stem completion tests. In both the free recall and recognition memory tests (tests of explicit memory), control subjects performed better than amnesic patients. Of course, poor (explicit) memory performance is a defining feature of the amnesic syndrome, so this finding was not surprising. However, the amnesic patients performed just as well as control subjects on the word fragment identification and word stem completion tests, the tests of implicit memory.

Based on the findings of Warrington and Weiskrantz, numerous researchers have argued that amnesic patients do not lack the ability to transfer verbal information from short-term to long-term memory, but instead have difficulty when asked to consciously retrieve such information. Evidence for dissociable forms of human memory is not limited to neurologically impaired populations alone: Dissociations between implicit and explicit forms of memory have also been demonstrated in normal human populations (reviewed in Roediger & McDermott, 1993). While some theorists have postulated distinct memory systems in the brain to explain these dissociations (e.g., Squire, 1987), others have proposed that these dissociations reflect different retrieval processes (e.g., Jacoby, 1983, 1988; Roediger, 1990a, 1990b). That debate aside, in simplest terms, implicit memory involves unconscious retrieval, while explicit memory involves conscious retrieval.

While explicit memory refers to the conscious recollection of past events, explicit learning can be characterized as the conscious, deliberate, analytical process used in discovering a rule (or sometimes a set of rules) that governs a set of stimuli. One classic example of an explicit learning task comes from Bruner, Goodnow, and Austin (1956), who applied a hypothesis-testing model to concept learning. In this study, subjects were presented with pictures that could be defined in terms of a number of relevant attributes, such as shape, color, and number of features. The pictures either were or were not members of a category, and membership in the category was based on a logical rule. For example, the stimuli in an experiment of this type could be pictures of flowers. A logical rule that would define membership in the category might be: “If a flower is red and has three petals, it is a member; otherwise it is not.” The experiment would progress as follows: Subjects were given a picture and had to decide whether the picture was a member of a category or not. After making a decision, subjects were told whether or not the picture belonged to the category and then were given another picture. Thus, for the first few stimuli, subjects simply had to guess about the membership of the stimulus. As the experiment progressed, however, subjects would eventually learn the rule that determines category membership (i.e., the concept) and begin using the rule to guide their decisions. Learning, in a typical experiment from Bruner et al.’s study, is explicit because it involves the systematic, analytic testing of specific hypotheses – that is, learning is explicit.

Researchers interested in how humans learn about complex stimulus domains, however, were not satisfied with the simplistic, well-defined, artificial structures used in early concept learning tasks like those of Bruner et. al. (1956). Clearly, simple logical rules are not representative of the complex rules that govern, for example, a natural language. One attempt to create a small-scale model of linguistic rules (which would end up being widely influential) was the finite-state (or artificial) grammar. An artificial grammar is an “algebra” for generating well-formed (“grammatical”) strings of symbols according to explicit rules (Chomsky & Miller, 1958). A diagram of an artificial grammar, used in an early experiment by Reber (1967), is shown in Figure 1. The numbered circles represent the states of the system, the arrows represent the permissible transitions from state to state, and the letters associated with each arrow represent the symbols generated by moving from one state to the next. Together, the system of states and transitions can be thought of as the “syntax” of the artificial language, and the particular letter set instantiated on the syntax can be thought of as the “vocabulary” of the artificial language. All well-formed strings in the artificial language are generated by starting at the input state and following any path through the diagram that ends at the output state. Although the complex grammar of English cannot be stated in such simple terms, these finite-state grammars were originally used as “miniature linguistic systems” to model language users (Chomsky & Miller, 1958; Miller & Chomsky, 1963).

One of the first important studies in artificial grammar research was done by George Miller (1958). Miller was inspired by the concept formation studies done by Bruner, Goodnow, and Austin (1956) and assumed that grammar learning could primarily involve the application of hypothesis formation and testing routines to experienced language productions. That is, the grammar of a language could be learned through an essentially explicit, analytical procedure analogous to the procedure used in concept formation tasks.

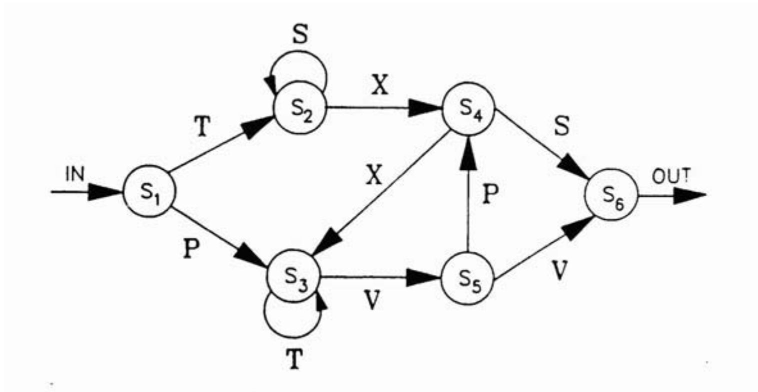


Figure 1. The artificial grammar originally used in Reber (1967). Grammatical letter strings are generated by moving through the grammar along the transitions from the input state (S_1) to the output state (S_6), and acquiring a letter with each transition.

To investigate this, Miller tested the hypothesis that familiarity with a system of grammatical rules might facilitate free recall. In his experiment, subjects were presented with a list of 9 letter strings, each ranging from 4 to 7 letters in length. One list contained strings that were formed with a random number table, while the other list contained strings formed by an artificial grammar (a grammar similar to, but less complex than the grammar shown in Figure 1). The letter strings were presented visually, one at a time, and subjects were given 5 seconds to memorize each string. Following the presentation of the entire list (all 9 strings), subjects were asked to recall, in any order, as many of the letter strings as possible. Subjects were given 10 trials with the same list of letter strings.

Miller found that the list of grammatical strings was learned much more quickly than the list of random strings: By the end of 10 trials, subjects recalling grammatical strings were able to recall on average almost all 9 letter strings, while subjects recalling random strings were able to recall on average only about 3 letter strings. Miller concluded that because letter strings generated by an artificial grammar were more redundant and carried less information than random letter strings, subjects were better at recoding the grammatical sequences into "chunks," thus reducing the difficulty of the task (Miller, 1958; also Miller, 1956). Miller's 1958 study was the only paper published on his artificial grammar learning research program, which Miller called "Project Grammarama" (Miller, 1967). Miller eventually abandoned Project Grammarama, feeling that the task of memorizing letter strings was terribly boring and that, ultimately, artificial grammar learning was conceptually limited. Miller believed there was insufficient common ground between his artificial grammar experiments and natural languages to make generalizations from one to the other (Miller, 1967).

Arthur Reber, however, believed otherwise. Reber became fascinated with the apparent ease with which subjects could classify letter strings as grammatical or non-grammatical after being exposed to a small subset of strings produced by an artificial grammar. In contrast to the explicit encoding explanation proposed by Miller (1958), where active, conscious, explicit processes are employed to discover grammatical rules, Reber had the intuition that passive, unconscious, "implicit" processes could be responsible for learning grammatical rules. Reber (1967) put his intuition to the test in a landmark study. His first experiment was a replication of Miller's (1958) study comparing free recall of letter strings generated randomly versus letter strings generated by an artificial grammar (see Figure 1). Reber replicated Miller's finding that grammatical letter strings were easier to learn and remember than random letter strings. However, Reber doubted Miller's explicit encoding hypothesis, saying, "It is not at all clear that an encoding system which is simple enough to discover yet efficient enough to facilitate learning even exists" (Reber, 1967, p. 859). Reber argued that if subjects were explicitly recoding individual symbols into larger "chunks," then they should have at least some verbalizable knowledge about the grammatical rules underlying the letter strings. However, when Reber informed subjects that the letter strings had been generated according to a complex system of rules and asked them to describe what they knew about the rules, subjects were unable to verbally express explicit knowledge about the rules of the grammar, even when asked specific questions about the letter strings (e.g., "What letter or letters may letter strings begin or end with?").

To further investigate whether the rules of the grammar were encoded in an explicit/analytic or implicit/non-analytic scheme, Reber conducted a second experiment, which would become the standard artificial grammar learning experiment for future implicit learning research. The experiment consisted of two phases, a learning phase and a testing phase. During the learning phase, subjects were asked to memorize 20 letter strings generated by an artificial grammar. The letter strings were presented in four sets of five strings, and the criterion for learning was two consecutive correct reproductions of a set. During the learning phase, subjects were not told that the letter strings followed any system of rules and were not informed that there would be a testing phase later in the experiment.

After completing the learning phase, subjects were told that the 20 letter strings they had just learned were formed according to a complex system of grammatical rules, though no information was given about the nature of the rules. Subjects were then presented with a series of novel letter strings and required to decide whether or not each of the strings followed the grammatical rules. Subjects were not given feedback about their decisions. Half of the novel letter strings were randomly generated and thus non-grammatical, while the other half of the letter strings were new grammatical strings that had not been presented during the learning phase. Reber found that subjects could classify the novel letter strings as grammatical or non-grammatical well above chance and, furthermore, subjects did not display any explicit, verbalizable knowledge about the rules of the grammar when questioned following the experiment. Reber concluded that subjects were able to learn about the structure underlying grammatical letter strings "without recourse to explicit strategies for responding or systems for recoding the stimuli" (Reber, 1967, p. 863). Subjects could then apply this implicit knowledge in a transfer task to discriminate grammatical and non-grammatical strings and, ultimately, could not verbally express their knowledge about any of the grammatical rules that were used to generate the sequences.

Before proceeding to review the developments in the field of implicit learning since Reber's (1967) seminal study, it is now useful to draw some preliminary distinctions between implicit and explicit learning and memory. The best method for distinguishing these related constructs is to consider the processes engaged in during encoding and retrieval (following Stadler & Roediger, 1998). This framework is presented in Table 1. As described earlier, implicit memory refers to facilitation in task performance without conscious recollection, while explicit memory refers to situations that require conscious recollection. Stated in terms of the encoding-retrieval distinction, explicit memory involves the conscious and deliberate retrieval of information, while implicit memory involves unconscious and non-deliberate retrieval. However, both implicit and explicit memory involve intentional, explicit encoding of information (e.g., studying a list of words). Like explicit memory, explicit learning involves analytic encoding (the discovery of a rule) and intentional retrieval (consciously applying the rule to new stimuli). Implicit learning, as of now defined only in terms of Reber's artificial grammar learning task, involves unintentional, non-analytic encoding of structural regularities in a stimulus environment. However, the testing phase in Reber's task makes explicit reference back to the original study phase (c.f., Schacter,

1987) by requiring subjects to make deliberate, explicit judgments about letter strings. Thus, the artificial grammar-learning task requires intentional retrieval of information. One cell in Table 1 remains empty: We have not yet described a situation where the encoding of information is largely unconscious and non-analytic, and the subsequent retrieval of that information is unintentional and non-deliberate.

Encoding	**Retrieval**	
	Conscious, Deliberate	Unconscious, Non-deliberate
Intentional, Analytic	Explicit Learning Explicit Memory	Implicit Memory
Unintentional, Non-analytic	Implicit Learning	

Table 1. Implicit and explicit learning and memory, characterized in terms of the processes engaged in during encoding and retrieval.

"What Is Learned" and "How It Is Learned"

Reber's research program on implicit learning went largely unnoticed until the late 1980's, when, for a number of reasons, implicit cognition received renewed interest. Several important papers relevant to implicit cognition were published around this time, concerning implicit memory (Schacter, 1987), dissociable memory systems (Squire, 1987), the cognitive unconscious (Kihlstrom, 1987), and a new laboratory technique, the serial reaction time task (SRT; Nissen & Bullemer, 1987), which would become just as popular as artificial grammar learning in the study of implicit learning (see Hsaio & Reber, 1998, for a review). Much of the renewed interest in implicit learning has been centered around two points of controversy, which will be called the "what is learned" and the "how it is learned" issues. Controversy surrounding these issues has emerged largely from Reber's claims that the process of implicit learning occurs largely outside of conscious awareness and that implicit learning yields a tacit knowledge base that is an abstract representation of the system of rules (Reber, 1989, 1993). Opponents of these views have argued that there is not sufficient evidence to support the claims that learning is unconscious and abstract (Dulany, Carlson, & Dewey, 1984, 1985; Perruchet & Pacteau, 1990, 1991). One extensive review and criticism of implicit learning concluded that dissociable forms of human learning do not exist (Shanks & St. John, 1994). This section reviews some of the literature surrounding the "what is learned" and "how it is learned" issues and suggests that much of this controversy stems from the process-purity assumption found throughout the implicit learning literature – that is, the assumption that a one-to-one mapping exists between a given task and an underlying process.

To address the "what is learned" issue, Reber (1969) conducted a follow up to his original study to assess the extent to which implicit learning is the acquisition of knowledge about the superficial physical form of the stimuli versus knowledge about the deeper, more abstract relations that underlie them. The study consisted of two sessions during which subjects memorized letter strings generated from an artificial grammar. Reber investigated the effects of changing certain aspects of the artificial grammar (the syntax, the vocabulary, or both) during the second session of the experiment. Subjects were assigned to one of four conditions. In the first control condition, neither the syntax nor the vocabulary were changed during the second session. In the second condition, the letter set (the vocabulary) remained the same, but the syntax, the system of rules governing the letter strings, was changed. In the third condition, the syntax remained the same, but the vocabulary was changed. Finally, in the fourth condition, both the syntax and the vocabulary were changed in the second session. The various manipulations had systematic effects on subjects' ability to memorize letter strings in the second session. When neither syntax nor vocabulary changed, subjects showed steady-state performance – that is, learning transferred from session one to session two. When both syntax and vocabulary changed, subjects showed a decrement in performance – learning did not transfer to session two. The most interesting finding, however, was that changing the physical form of the stimuli (changing the vocabulary) did not have an adverse effect on the transfer of learning – learning could be transferred across letter sets. On the contrary, changing the system of rules governing the stimuli (changing the syntax) produced a decrement in performance in session two. These findings support the argument that the implicit knowledge acquired in the first session was abstract and did not depend on the physical form of the stimuli. Reber concluded that implicit learning was the abstraction of the deep structure, the syntactic system of rules that governed the stimuli rather than simply the learning of the explicit symbols of the surface patterns.

Several other studies have also employed transfer of learning paradigms to argue that implicit learning is a process of rule abstraction that is not tied to the physical form of the stimulus. A study by Mathews, Buss, Stanley, Blanchard-Fields, Cho, and Druhan (1989) replicated and extended Reber's original finding that implicit learning can transfer across letter sets. Mathews et al. tested subjects in an artificial grammar-learning task over a period of four weeks. Subjects who received a new letter set each week of the experiment, without any change to the underlying syntax, performed just as well as subjects who worked with the same letter set throughout the four-week period. Mathews et al. described this transfer across vocabularies as occurring "immediately and automatically without any conscious translation process."

Two other artificial grammar learning studies demonstrating transfer across sensory modalities are also worth mentioning here. Altmann, Dienes, and Goode (1995) presented subjects with sequences of auditory tones during the learning phase and then had subjects make grammaticality judgments about sequences of visually presented letters during the testing phase. They found that exposure to the grammar in the auditory modality improved classification of novel stimuli in the visual modality. In a more complex series of experiments, Manza and Reber (1997) compared cross-modal transfer conditions (from visual learning to auditory testing, and vice versa) with within-modality conditions. Manza and Reber found no differences between subjects in these four conditions in their ability to classify strings during testing. Together, these two transfer of learning experiments provide support for the "abstractionist viewpoint," that is, that implicit learning involves the abstraction of deep structure out of a stimulus environment.

Critics of the abstractionist view have developed innovative ways of demonstrating that fragmentary knowledge about memorized letter strings can explain classification performance in an artificial grammar-learning task. In a study by Dulany, Carlson, and Dewey (1984), subjects were given an artificial grammar learning task, following the standard Reber procedure, except that during the testing phase subjects were asked to specify the feature or features of each test string that led them to classify it as they did. That is, for items that were judged grammatical, subjects were asked to underline the letter or letters they felt made the item grammatical; for items they felt were not grammatical, subjects were asked to cross out the letters that rendered the string non-grammatical. Consistent patterns were found between subjects' grammaticality decisions and the portions of the strings they indicated as guiding those decisions. Based on these findings, Dulany et al. argued that subjects were not making their decisions based on any abstract representation of rules, but instead were basing decisions on knowledge about particular fragments in the test strings.

In a related study, Perruchet and Pacteau (1990) also claimed that subjects did not induce abstract representations of the artificial grammar but, instead, established fragmentary representations based on patterns of bigrams. To support their argument, Perruchet and Pacteau used two different learning techniques. One group of subjects studied a list of whole letter strings, while another group of subjects studied a list of frequently occurring permissible bigrams. Perruchet and Pacteau found that both groups were able to classify grammatical and non-grammatical test strings above chance and concluded that Reber's abstractionist view was unfounded. If subjects could classify strings better than chance after being exposed only to permissible bigrams, then it was likely that subjects' resulting representations were fragmentary as well.

Finally, Brooks and Vokey (1991; Vokey & Brooks, 1992) presented what is perhaps a compromise in this debate. They argued that grammaticality judgments can be made reliably without having a purely abstract representation; instead, all that is needed is an "instantiated memory" consisting of specific items and a decision making process based on a similarity metric. Brooks and Vokey examined performance in an artificial grammar-learning task while factoring in the physical and structural similarity of the test items. In the testing phase, subjects were presented with four types of strings: "close-grammatical" strings (i.e., ones that were physically close to a string memorized during learning and conformed to the rules of the grammar), "far-grammatical" strings (i.e., ones that differed by two or more elements from a memorized string and conformed to the rules of the grammar), "close-non-grammatical" strings (i.e., ones that were physically close to a learning string but violated the rules of the grammar), and "far-non-grammatical" strings (i.e., ones that were physically remote from learning strings and violated the rules of the grammar). Brooks and Vokey found that roughly half of the variance in subjects' performance could be attributed to each underlying factor, physical and structural similarity. Based on these results, Brooks and Vokey concluded that subjects were using both abstract and concrete representations to classify test strings.

To address the "how it is learned" issue, Reber (1976) manipulated the instructional set given to subjects in his artificial grammar learning task. One group of subjects was given a "neutral" instructional set, replicating the procedure of his original study (Reber, 1967). A second group of subjects was informed that a rule system governing the stimuli existed. These subjects were encouraged to search for the structure in the stimuli. Nothing was explicitly told to this group about the nature of the structure underlying the letter strings. Both groups were given the same learning phase, during which they memorized strings produced by an artificial grammar, and the same testing phase, during which they decided whether novel letter strings were or were not grammatical.

Remarkably, the explicitly instructed group of subjects performed worse than the group given neutral instructions in all aspects of the experiment: They took longer to memorize letter strings during the learning phase, they were worse at discriminating grammatical from non-grammatical strings during the testing phase, and they showed evidence of having induced rules that were not representative of the grammar in use. These results suggest that under these conditions explicit processing of complex materials has a disadvantage relative to implicit processing. Reber (1976) argued that, in this particular task, the explicit instructions actually had an interference effect: Subjects were being encouraged to search for rules that, given the nature and the complexity of the artificial grammar, they were unlikely to find in the surface patterns. In Reber's own words, "The simplest conclusion here seems to be the right one: Looking for rules won't work if you cannot find them" (Reber, 1993, p. 48).

Clearly, there is a difference between simply informing subjects that the stimulus materials have some underlying structure and giving subjects precise information about the nature of that structure. A follow-up study by Reber, Kassin, Lewis, and Cantor (1980) attempted to address this issue directly. Subjects were presented with the actual schematic structure of the artificial grammar (i.e., Figure 1) and were given a seven-minute "course" on how the grammar works to generate letter strings. At the same time, subjects were shown 20 strings produced by the grammar, as was done in the standard artificial grammar learning task. Reber et al. manipulated the time when explicit training was introduced: One group was given the grammar at the beginning of the learning phase, a second group received it halfway through the learning phase, and the third group was given explicit training only after seeing all of the exemplars in the learning phase.

The key finding from this study was that the earlier subjects were given explicit training, the better they performed on the classification test. Reber et al. suggested that providing explicit training at the outset of the learning phase focuses the subjects' attention on the relevant structural relationships in the letter strings. The explicit training did not teach the full grammar to the subjects, Reber et al. stressed, but rather oriented subjects to the relevant dimensions and relationships in the letter strings.

The critics of the abstractionist view are largely the critics of the unconscious view as well, and for similar reasons. Dulany et al. (1984) argued that subjects' classification decisions were based on their conscious awareness of the presence or absence of key fragments in the test strings. Thus, subjects never derive any information about the actual grammar but instead develop idiosyncratic, correlated grammars via conscious, analytic processes. Perruchet and Pacteau (1990) made a similar argument, that subjects consciously encode frequently occurring fragments of strings presented in the learning phase and then apply this knowledge during the testing phase. Furthermore, Perruchet and Pacteau argued that the actual learning involved in making grammaticality judgments occurred during the testing phase, not in the learning phase, because at that time subjects are aware of the existence of the grammar while they are making classification decisions. Thus, conscious, fragmentary knowledge is strategically applied in the artificial grammar-learning task.

In addition to these criticisms, perhaps the most influential argument against unconscious learning is theoretical in nature. Suppose that implicit and explicit (unconscious and conscious) processes can be represented as two variables, α and β respectively. Implicit learning is said to have occurred in the absence of conscious awareness. That is, the circumstances for establishing the existence of implicit learning are $\alpha > 0$ and $\beta = 0$. Proving the latter, that $\beta = 0$, is highly problematic. In response to Reber's (1989) review of implicit learning, Brody (1989) argued that nowhere in the entire body of implicit learning research have the "proper" tests been run to determine whether subjects had any conscious knowledge of the rules or underlying system. That is, no study of implicit learning has ever established that $\beta = 0$.

In their recent critique of implicit learning research, Shanks and St. John (1994) expand upon the idea proposed by Brody (1989) that no study has provided sufficient evidence that implicit learning is unconscious. Specifically, Shanks and St. John pointed out two problems inherent in equating consciousness with verbal report. The first, what they called the "Sensitivity Criterion" is that a test of verbal report may not be sensitive to all of the conscious knowledge the participant has available. For example, how does the experimenter know he has asked the right questions and pressed the subject hard enough? The second problem with equating consciousness with verbal report is the potential violation of the "Information Criterion." This is the notion that a subject's task performance may depend on information I, but in the test of awareness the experimenter is mistakenly looking for information I*. Shanks and St. John cite the findings of Perruchet and Pacteau as an example of this discrepancy. In their study, the information needed to complete the task (I) was information about permissible bigrams, while the information requested in a standard verbal report concerns syntactic rules (I*). Ultimately, Shanks and St. John concluded that there has not been sufficient evidence demonstrating the existence of unconscious learning and that dissociable forms of human learning do not exist.

In response to these criticisms, Reber (1993) has argued that establishing $\alpha > 0$ and $\beta = 0$ is not necessary for demonstrating implicit learning. Instead, he suggests the relationship can be stated as a simple inequality, $\alpha > \beta$, for a circumstance where implicit knowledge is at work. In other words, when explicit, conscious knowledge is not sufficient to explain performance on a given task in some domain, then implicit, unconscious knowledge must account for this discrepancy. Reber argued that, "The key to uncovering the cognitive unconscious will be found in measures of mental content held consciously (β) that yield lower estimates when compared to those made of the mental content held outside the purview of consciousness (α)" (Reber, 1993, pp. 8-9). The only ontological stance that will be successful in studying unconscious, implicit mental processes holds that the unconscious is the "default condition", a stance Reber refers to as "the primacy of the implicit" (Reber, 1990, 1993). This position will be expanded upon in the following section describing Reber's evolutionary model of the cognitive unconscious.

The controversies surrounding the issues of "what is learned" and "how it is learned" still remain unresolved. The only conclusion that will be drawn here is that throughout the recent history of research on implicit learning, researchers have consistently made the process-purity assumption, the assumption that a one-to-one mapping between a given task and an underlying cognitive process can exist. Clearly, this dubious assumption is at the root of criticisms arguing that conscious, explicit knowledge is never entirely absent, that you can never get $\beta = 0$.

Reber's Evolutionary Model of the Cognitive Unconscious

Reber's motivation to develop a theory of the cognitive unconscious from an evolutionary perspective stemmed from his desire to place contemporary cognitive science within a solid biological framework (Reber, 1992a, 1992b, 1993). Reber was largely dissatisfied with information processing models of cognition that postulated metaphorical boxes and buffers and with nativistic theories of mind (e.g., Fodor, 1983), such as the proposal of an immensely complex and highly specific language acquisition device (e.g., Chomsky, 1986; Pinker, 1994). According to Reber, these proposals lack biological coherence and ecological validity. Instead, a better proposal would be that distinct types of cognitive functions (implicit and explicit, unconscious and conscious) evolved at different times for different reasons. Reber's evolutionary model of the cognitive unconscious is based on four principles from evolutionary biology. The model is useful in the specific predictions it makes concerning dissociations between implicit and explicit cognitive processes. The principles are as follows:

1. The Principle of Success: Forms that have proven their adaptive value become the foundation for later forms.

2. The Principle of Conservation: Successful forms become fixed and begin to serve as foundations for developing forms– thus their core features are unlikely to be substantially modified over time.

3. The Principle of Stability: Early appearing, successful forms tend to be relatively invariant.

4. The Principle of Commonality: Evolutionarily earlier forms will be displayed across species.

Consciousness, Reber argues, was a relatively late arrival on the evolutionary scene and was preceded by sophisticated cognitive functions that operated automatically without its benefit – that is, unconscious/implicit processes. Since implicit cognitive processes are evolutionarily older, Reber emphasizes what he calls the "primacy of the implicit" (Reber, 1990). What he means is that unconscious, implicit processes are the "default mode" and are the foundation upon which emerging conscious, explicit operations have been laid. Based on this evolutionary model, Reber makes several specific predictions about the properties of implicit processes that differentiate them from explicit processes:

1. Robustness: Implicit processes should remain intact in the face of disorders that affect explicit processes.

2. Commonality: The underlying processes of implicit learning should show cross-species commonality.

3. Age independence: Implicit learning should show few effects of age and developmental level compared with explicit learning.

4. Low variability and IQ independence: Implicit learning should show little or no individual differences and also should show little concordance with measures of "intelligence" assessed by standard psychometric instruments.

The first two predictions will not be discussed at length here. As mentioned earlier, implicit memory appears to be left intact following neurological damage (Warrington & Weiskrantz, 1970); evidence has also been found showing preserved implicit learning in the face of various neurological impairments (Knopman & Nissen, 1987; Knowlton & Squire, 1994, 1996; Nissen & Bullemer, 1987). The prediction that implicit learning processes are present in multiple species is largely philosophical in nature and has been discussed in depth elsewhere (Reber, 1993, 1997b). The third and fourth predictions, however, are specific predictions about individual differences in implicit learning that are relevant to the present study. Surprisingly, only a few studies in the literature have specifically addressed the issue of individual differences in implicit learning.

The first study to investigate the relationships among implicit and explicit learning and IQ was conducted by Reber, Walkenfeld, and Hernstadt (1991). Reber et al. compared performance on the standard artificial grammar learning task with performance on an explicit problem-solving task and four subtests from the Wechsler Adult Intelligence Scale-Revised (WAIS-R). The problem-solving task involved predicting the next item in a sequence of letters. For example, given the sequence ABCBCDCDE__, the next letter would be D, since the sequence is broken into chunks of three, each chunk beginning with the next letter in the alphabet after the letter that initiated the previous chunk (ABC-BCD-CDE-__).

Performance on the implicit and explicit tasks showed the predicted pattern of correlations. The correlation between performance on the explicit problem-solving task and IQ was significant ($r = +.69$, $p < .01$); the correlation between the artificial grammar learning task and IQ was not significant ($r = +.25$, $p > .05$); and the implicit and explicit tasks did not correlate significantly with each other ($r = +.32$, $p > .10$). Reber et al. concluded that these findings supported the proposal that implicit tasks are fundamentally different from explicit tasks and that these differences are best viewed within an evolutionary framework (Reber, 1993).

A similar study by Mayberry, Taylor, and O'Brien-Malone (1995) compared performance on implicit and explicit tasks with IQ in school-aged children, Grades 1-2 and 6-7. Mayberry et al. found that performance on an implicit learning task was not related to IQ (*r*'s = +.02 and +.04 for children in Grades 1-2 and 6-7, respectively), but that performance on an explicit task was (*r*'s = +.37 and +.56). Also, the degree of verbalizable knowledge about the implicit test was not related to performance on the implicit test (*r*'s = +.05 in both groups), while the degree of verbalizable knowledge about the explicit test did correlate with performance on the explicit test (*r*'s = +.47 and +.80). These findings provide support for the earlier conclusions made by Reber et al. (1991) that implicit learning does not co-vary with either explicit learning or traditional psychometric measures of intelligence.

McGeorge, Crawford, and Kelly (1997) replicated and extended the findings of Reber et al. (1991). Subjects in this experiment ranged from age 18 to 77 years, allowing a more thorough investigation of the development of implicit and explicit functions across the life span. Subjects also completed a full-scale measure of IQ (the WAIS-R), and thus a factor analysis could be performed to better determine what components of the IQ score were most related to the implicit and explicit tests.

McGeorge et al. found that while the correlation between performance on the implicit test and overall IQ was not significant ($r = +.12$), there was a small but significant loading on the Perceptual Organization factor ($r = +.19$), a factor thought to be associated with general fluid intelligence. In contrast, the explicit test showed strong correlations with overall IQ ($r = +.67$) and with both the Perceptual Organization and Attention-Concentration factors (*r*'s = +.65 and +.53, respectively). Finally, McGeorge et al. found that while there were no differences in performance on the implicit test with increasing age, performance on the explicit test did decrease with age. Thus, the findings reported by McGeorge et al. replicate and extend those of earlier studies comparing implicit and explicit tests with psychometric intelligence and, additionally, are relevant to Reber's (1993) prediction that implicit cognitive functions will be age independent.

Almost all of the discussions concerning individual differences in implicit learning conclude that they do not exist, and this follows from the theoretical arguments proposed by Reber (1992a, 1992b, 1993, 1997a). However, the empirical studies described above demonstrate only that implicit learning does not co-vary with explicit measures of learning, which is not evidence that individual differences in implicit learning do not exist. Furthermore, finding a relationship between explicit tests and IQ and no relationship between implicit tests and IQ, is a foregone conclusion: Psychometric tests of intelligence are designed to measure the same construct thought to underlie an explicit test, not to measure implicit cognitive processes. None of the studies reviewed earlier have focused specifically on finding variability in performance on an implicit learning task. Indeed, the actual variability among individual performance in implicit tests has largely been ignored in these studies. Only recently has Reber softened his stance concerning the existence of individual differences in implicit learning, noting that the studies by both Reber et al. (1991) and Mayberry et al. (1995) did find some individual variation in success on the implicit task, allowing for the possibility that individual differences in implicit learning do exist (Reber & Allen, 2000). However, these individual differences remain unexplained and have not been the focus of any systematic research effort.

To study individual differences, Reber suggests, "If we are to come to some conclusions about inter-individual differences in implicit learning, we need to examine individual performances on implicit and explicit tasks which use a common metric" (Reber & Allen, 2000, p. 241). The argument here is for a task dissociation: A common metric is needed so that a direct comparison can be made between an implicit task and an explicit task. However, this appears be the same strategy used in Reber et al. (1991), Mayberry et al. (1995), and McGeorge et al. (1997), so there is no reason to believe that yet another comparison between an implicit task and an explicit task will provide any further insights into the nature of individual differences in implicit learning. An alternative approach to measuring individual differences in implicit learning would be to first separate the relative contributions of implicit and explicit processes to performance on a specific task, and then attempt to account for variability in the implicit process with some other measure. Based on this reasoning, a new method for measuring individual differences in implicit learning is proposed and developed in the following section.

Measuring Individual Differences in Implicit Learning

As mentioned earlier, almost all of the research on implicit learning has made the process-purity assumption, the assumption that a one-to-one mapping between a given task and an underlying cognitive process exists. Specifically, the theory of implicit learning originally assumed that learning in an artificial grammar-learning task is unconscious (Reber, 1967, 1969, 1976). Critics of implicit learning then argued that learning in an artificial grammar-learning task is accompanied by conscious awareness (Dulany et al., 1984; Perruchet & Pacteau, 1990). A revised stance, proposed by Reber (1993, 1997a), is that learning in an artificial grammar learning task is largely unconscious, but not entirely so. This revised view dismisses the process-purity assumption, since both conscious and unconscious processes can function in a given task. However, simply observing that there is an inequality between the amount of information available for conscious expression and that available to the unconscious does not provide a means for separating the relative contributions of implicit and explicit processes to performance on a given task.

Over the last decade, Jacoby (1991; Jacoby, Toth, & Yonelinas, 1993) has developed a process dissociation procedure in attempt to solve a similar problem, separating automatic and intentional (controlled) influences in memory. Several assumptions from Jacoby's process dissociation framework can also be applied to the study of implicit and explicit influences on learning. The first critical assumption is the rejection of process-purity: A given task does not have to measure only one underlying process but, instead, may simultaneously engage both implicit and explicit processes to varying degrees. Thus, task dissociations (e.g., comparing classification performance with verbalizable knowledge) do not necessarily reflect process dissociations. Second, the processes driving performance on a given task can work in concert to facilitate performance, in opposition to worsen performance, or may have neutral effects on each other. Borrowing an example from Jacoby (1991), automatic (A) and intentional (I) processes can work together to facilitate memory (A + I) or can act in opposition to interfere with memory (A – I). Third, arranging tasks so that performance in a facilitation paradigm can be compared with performance in an interference paradigm will allow us to look for process dissociations, rather than look for task dissociations. Finally, Jacoby argues that a measure of the implicit, unconscious influences on learning obtained within a process dissociation framework will be less susceptible to the prevailing methodological criticisms in the implicit learning literature, largely concerning task dissociations, and may prove to be insightful about the nature of any possible individual differences in implicit learning.

A recent study by Destrebecqz and Cleeremans (2001) applied the process dissociation procedure to an implicit sequence-learning task, the serial reaction time (SRT) task, developed by Nissen (Nissen & Bullemer, 1987). In a typical SRT task, subjects are presented with a series of sequentially structured stimuli, usually on a row of four lights. On each trial, subjects see a light in one of the four locations "light up" and are asked to press as fast and as accurately as possible a key corresponding to the location of the light. For some subjects, the sequence of lights is generated randomly, while for others the sequence of lights follows a repeating pattern. The typical finding in SRT experiments is that reaction times (RTs) tend to remain constant across trials in the random lights group, while RTs progressively decrease across trials in the repeating pattern group. Furthermore, RTs will increase dramatically if the repeating pattern is abruptly modified. These findings suggest that subjects are able to learn to respond to the repeating pattern of lights, even though subjects fail to demonstrate verbalizable knowledge about the pattern (Nissen & Bullemer, 1987; Willingham, Nissen, & Bullemer, 1989).

Destrebecqz and Cleeremans (2001) examined the contribution of explicit knowledge in an SRT task by using a free generation task and a recognition memory task. After being trained in a standard SRT task, subjects were informed that the stimuli they had seen earlier had actually followed a repeating pattern. Subjects were then asked to freely generate several trials that "resembled the training sequence as much as possible." This generation task was performed under "inclusion" instructions, so that presumably implicit and explicit knowledge bases would act in concert to facilitate performance (E + I). After performing the generation task in the facilitation condition, subjects were asked to freely generate several sequences under "exclusion" instructions – that is, they were told to avoid reproducing the sequential regularities of the original training sequence. In this condition, an implicit knowledge base (if implicit learning had occurred) would interfere with performance on the generation task (E – I). Following both generation tasks, subjects were given a recognition memory task that asked them to decide whether fragments of sequences were part of the original training sequence. Subjects provided a rating, on a six-point scale, of how confident they were that the fragment had occurred in the training sequence.

Destrebecqz and Cleeremans (2001) found that subjects were able to produce some explicit knowledge of the training sequence in the free generation task under inclusion instructions. However, when subjects were asked to deliberately exclude sequential regularities from the training sequence, they were unable to avoid projecting their implicit knowledge in the sequences they generated. In the recognition memory test, subjects were mostly unable to discriminate old and new fragments of the training sequence. Destrebecqz and Cleeremans concluded that in attempting to separate the relative contributions of implicit and explicit processes, their findings provided evidence that sequence learning can be unconscious.

The application of the process dissociation procedure used in the present study is similar to that of Destrebecqz and Cleeremans (2001). The task, however, was designed to better quantify implicit knowledge using a measure that is much more sensitive to variability among individuals. Looking back on the original studies of artificial grammar learning, Miller (1958) and Reber (1967) both showed that subjects were better at memorizing letter strings governed by a grammar than at memorizing random letter strings. One question that was not addressed in these early studies, and has not been addressed since that time, is whether having learned an artificial grammar will improve later memory for strings generated by that same grammar. Simply put: Can implicit knowledge about an artificial grammar improve memory span for sequences governed by that grammar?

A significant early experimental psychology textbook describes memory span as folows: "The concept of span, derived from the span of the hand, conveys the idea of width of grasp. How much can be spanned or grasped at once?" (Woodworth, 1938). Memory span as a dependent measure can be used as an alternative to classification performance and lends itself nicely to the process dissociation procedure. Immediate serial recall is generally accepted as a measure of explicit memory that involves conscious encoding and retrieval of a list of items. The goal of this study was to design a new task where memory span performance, a largely explicit cognitive process, would be facilitated by prior implicitly learned knowledge (E + I). Comparing performance on this task with performance on an equivalent memory span task where there is no help from prior implicitly learned knowledge (E + 0) could then yield a measure of the relative contribution of implicit learning (I).

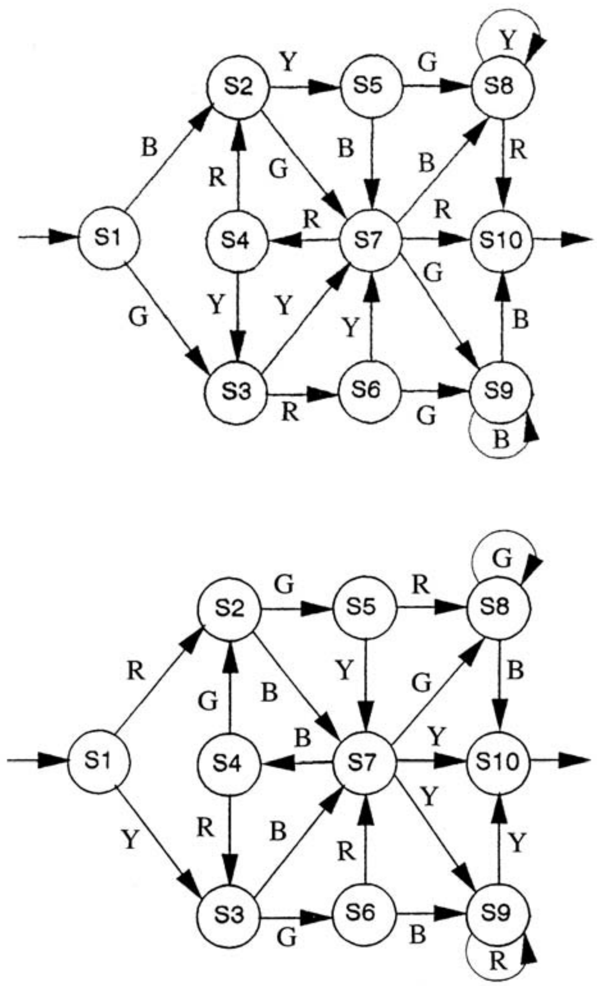


Figure 2. The two artificial grammars used to generate sequences in this experiment, adapted from Brooks & Vokey (1991). The grammars share the same syntax and vocabulary, but differ in their syntax-vocabulary arrangements.

The task used in the present experiment was relatively simple and straightforward. Subjects were presented with a sequence of colors and were asked to immediately reproduce the sequence, using a custom designed response box. Subjects performed this same task during an "acquisition phase" and a "test phase." The sequences presented to the subjects during the acquisition phase were generated by an artificial grammar (see Figure 2). By reproducing sequences during the acquisition phase, subjects were given exposure to the underlying grammar. During the test phase, half of the sequences presented were new sequences that came from the same grammar that the subject had been trained on. The other half of the sequences presented during the test phase were new sequences generated by a different grammar, to which the subject had not been given any prior exposure. Thus, the ability to remember and reproduce sequences that came from the "Trained" grammar would be facilitated by implicitly learned knowledge (Trained = E + I). However, performance on "Not Trained" sequences would not be facilitated by any implicit knowledge (Not Trained = E + 0).

In carrying out this task, however, it is possible that subjects could consciously and explicitly remember sequences that occurred during the acquisition phase and then use that explicit knowledge to improve their memory span during the testing phase. In other words, it is possible that remembering and reproducing sequences from the "Trained" grammar is entirely due to explicit knowledge (Trained = E + E). To account for this possibility, an additional task was included that would place implicit and explicit knowledge bases in opposition.

After completing the acquisition and test phases, subjects were given a recognition memory test that consisted of several types of sequences. "Old" sequences, selected from the first part of the experiment, came from either the "Trained" grammar or the "Not Trained" grammar. Three types of "New" sequences were used: Ones that came from the "Trained" grammar, ones that came from the "Not Trained" grammar, and ones that were randomly generated and did not conform to the rules of either grammar. If subjects had developed an explicit knowledge base during the acquisition phase, then they should be best at discriminating between old and new sequences that came from the Trained grammar (because Recognition = E + E). However, if implicit learning occurred during the acquisition phase of the experiment, then subjects should be worst at discriminating between old and new sequences, because implicit knowledge about the grammar and explicit memory for actual sequences would be set in opposition (Recognition = E – I).

The recognition memory test was also specifically designed to assess subjects' judgments about their phenomenological experience while recognizing grammatical sequences, using a procedure developed by Tulving (1985). During the recognition memory test, subjects were presented with a sequence and asked to judge whether the sequence was "old" or "new". If a sequence was judged old, then subjects were asked to further distinguish between two states of awareness: remembering and knowing. A "remember" experience is defined as one in which the subject can mentally relive the original occurrence of the sequence. A "know" judgment is made when subjects are confident that the sequence had occurred before but are unable to re-experience its occurrence. There is a sizeable literature on remember and know judgments (see Rajaram & Roediger, 1997) providing evidence that remember-know judgments do not simply reflect states of confidence, since several variables can differentially affect remember-know judgments and confidence ratings in these tasks (Rajaram, 1993).

The sequences of colors were presented to subjects using the Simon memory game (see Figure 3), a method of measuring immediate memory span that has been developed in our laboratory (Cleary, Pisoni, & Geers, 2001). In previous research, we have used the memory game to measure immediate serial recall in deaf children with cochlear implants, in normal-hearing children, and in normal-hearing adults (Pisoni & Cleary, 2002). The memory game allows stimuli to be presented to subjects using three different stimulus presentation formats. One group of subjects was presented with visual sequences on the memory game (sequences of colored lights); the second group was presented with auditory sequences (sequences of spoken color words); and the third group was presented with auditory + visual sequences (sequences of colored lights and spoken color words presented simultaneously). After an entire sequence was presented, subjects simply reproduced the pattern by pressing the response buttons on the memory game response box.

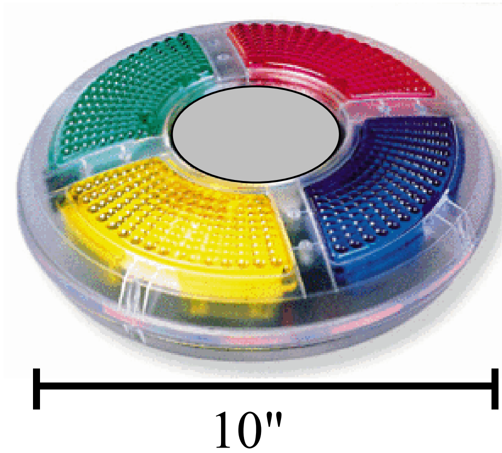


Figure 3. The Simon memory game. Subjects reproduce sequences of colors, presented in auditory, visual, or auditory + visual formats, by pressing the colored response buttons.

The specific predictions in this study were, first, that subjects would have larger memory spans for sequences that came from the grammar they were trained on versus sequences that came from the grammar on which they were not trained. Second, the particular grammar learned by the subject should not make any difference in performance: A comparable learning effect should be found for subjects trained on Grammar A and for subjects trained on Grammar B. Third, subjects should be worse at discriminating old-new sequences from the "Trained" grammar than at discriminating old-new sequences from the "Not Trained" grammar. In other words, subjects should be more likely to falsely recognize, and falsely remember, sequences from the "Trained" grammar, compared with sequences from the "Not Trained" grammar or randomly generated sequences. Such a finding would indicate that implicitly learned knowledge about the "Trained" grammar interfered with the ability to correctly reject a new grammatical sequence. Lastly, individual differences in the size of the learning effect (i.e., variation in the degree of implicit learning) should co-vary with a standard, unrelated explicit measure of immediate memory span, assessed using a traditional digit span task.

It may be helpful to place the current investigation of individual differences in implicit learning within a somewhat larger framework that deals directly with an important clinical problem. The Simon memory game was originally developed as a method of measuring memory span in congenitally deaf children who use cochlear implants, a device that provides access to sound by delivering electrically-coded signals to the auditory nervous system (Miyamoto, 1995). A wide array of individual differences in outcome measures of speech perception and overall language development has been observed in this population of children (Pisoni, Cleary, Geers, & Tobey, 2000; Pisoni, Svirsky, Kirk, & Miyamoto, 1997). At the present time, there is no clear explanation of why some children appear to do well at acquiring language with their cochlear implants while other children do not. A substantial portion of the research in our laboratory has been devoted to investigating the nature of individual differences in the language development of deaf children who use cochlear implants (Cleary, Pisoni, & Geers, 2001; Pisoni, 1999, 2000; Pisoni & Cleary, 2002; Pisoni & Geers, 2000; Pisoni et al., 1997; Pisoni et al., 2000).

A central thesis in this research has been that the operation of higher-level cognitive processes may ultimately be responsible for the variability in language development observed in this population of children. It has been argued elsewhere that immediate memory span and other measures of cognitive processes such as encoding, storage, and retrieval of information form the foundation for language abilities (Gupta & MacWhinney, 1997). However, traditional methods of measuring memory span involve the verbal reproduction of a presented list of items, which could confound results obtained from this population of children, because of articulation problems. Some very young children with cochlear implants have difficulty producing intelligible speech, due to differences in speaking rate and fluency of articulation. Thus, the Simon memory game was developed in order to measure memory span without requiring an explicit verbal response.

While the endeavor to account for individual differences in the language abilities of deaf children with cochlear implants has been successful (Pisoni et al., 2000), none of the experimental techniques in this research program have been developed specifically to measure individual differences in implicit cognitive processes. Indeed, most researchers believe either that individual differences in implicit learning do not exist or that they are not a significant research issue (e.g., Reber, 1997a). Many arguments have been made that the emergence of language abilities is due largely to implicit mechanisms (Gupta & Cohen, in press; Kirsner, 1994; Reber, 1967, 1989, 1993; Winter & Reber, 1994). The argument follows, then, that individual differences in language abilities might be a function of individual differences in underlying implicit processes. Thus, investigating the possibility that individual differences in implicit learning are responsible for some portion of the observed individual differences in language abilities (in any population – but in particular, among deaf children who use cochlear implants) is potentially a highly significant research issue, particularly for a research program interested in understanding the underlying basis of individual differences in order to develop new intervention strategies to improve the language abilities in those children who show relatively poor development (Pisoni et al., 2000).

Method

Subjects

One hundred and twenty Indiana University undergraduates, ages 18 to 24, participated in this study for partial fulfillment of course requirements for introductory psychology. All participants were native speakers of English with no speech or hearing disorders and had normal or corrected-to-normal vision at the time of testing.

Materials

Digit Span. Tokens of the 10 spoken digits ("0" to "9") obtained from the Texas Instruments 46-Word (TI46) Speaker-Dependent Isolated Word Corpus (Texas Instruments, 1991) were used for the auditory digit span task. Auditory stimuli were presented over high-quality headphones at 75 dB SPL. Subjects made their responses by writing in prepared answer booklets at the end of each trial. After recording their responses, subjects initiated the next trial by pressing the "Enter" key on the keyboard.

Simon Memory Game. Auditory tokens of the four color words ("red," "yellow," "green," and "blue") were recorded by one male speaker of American English. The memory game response box was modeled after the commercial product "Simon" by Milton Bradley (see Figure 3), which consists of four colored, back-lit response buttons. Subjects made their responses to auditory, visual, or auditory-plus-visual stimuli by pressing the response buttons on the memory game (Cleary, Pisoni, & Geers, 2001).

Two artificial grammars (referred to as Grammar A and Grammar B) were used to generate grammatical sequences. The grammars were adapted from Brooks and Vokey (1991) and are shown in Figure 2. These particular grammars were chosen because they could generate a greater number of short grammatical sequences than other frequently used grammars (e.g., Figure 1; Reber, 1967). Because Grammar A and Grammar B share the same syntax and the same vocabulary, the sequences they generate should be equally complex and equally difficult to remember and reproduce. Each grammar consists of a vocabulary (the colors "red," "yellow," "green," and "blue," represented as R, Y, G, and B in the diagram), and a syntax, which is the set of states (represented as circles) and the transitions between those states (represented as arrows). The generation of a grammatical sequence begins by entering the grammar at state S_1. Each transition from any state S_i to any state S_j produces an item in the sequence. The sequence ends when state S_{10} is reached. The sequence of colors produced depends upon the path taken through the state diagram of the grammar.

The number of possible sequences at a given length that each grammar can generate was determined for lengths 4 through 10. Additionally, the number of possible random sequences at each length was determined. These values, shown in Table 2, illustrate how the artificial grammars used in this study function to constrain the number of possible acceptable sequences. For example, at length 7, 30 sequences out of a possible 16,384 random sequences (0.18%) conform to each of the grammars.

Grammatical sequences used in the experiment were selected pseudo-randomly from the set of all possible grammatical sequences from length 4 to length 10. No sequence with more than three consecutive repetitions of a given color was used. Sequences were selected for the acquisition phases so that each branch of the grammar was represented. Sequences were selected for the test phase to ensure that no sequence occurred in both acquisition and testing.

List Length	No. of Grammatical Sequences	No. of Random Sequences
4	8	256
5	10	1024
6	14	4096
7	30	16,384
8	46	65,536
9	58	262,144
10	98	1,048,576

Table 2. The number of possible grammatical sequences that Grammar A and Grammar B can generate at each list length, from length 4 to length 10, compared with the total number of possible random sequences at each length.

An additional set of non-grammatical sequences, lengths 5 through 7, was randomly generated for use in the recognition memory test. All random sequences were checked to ensure that they did not conform to the rules of either grammar. These random sequences were allowed to begin with any of the four colors in the vocabulary. However, an equal number of random sequences beginning with each color was used in the recognition memory test.

Procedure

Subjects were tested in groups of three or fewer in a sound attenuated testing room. All subjects first completed a digit span task. Then they received the acquisition phase, test phase and a final recognition memory test on the Simon reproduction task.

Digit Span. Subjects were presented with a list of digits over headphones. Once the entire list had been presented, subjects wrote down as many digits from the list as they could remember, in the order in which they were originally presented. The lists of digits began at length 4 and increased to length 10, with two lists presented at each length, for a total of 14 trials.

Acquisition Phase. Using the Simon memory game, subjects were presented with a sequence of colors and were simply asked to reproduce the sequence by pressing the response buttons. Color sequences were presented either auditorily (a sequence of spoken color words), visually (a visual-spatial sequence of colored lights), or audiovisually (a visual-spatial sequence of colored lights and the same sequence of spoken color words, presented simultaneously). Stimulus presentation format was a between-subjects factor. Subjects were only exposed to sequences that came from one grammar (A or B). The grammar used during acquisition was counterbalanced across subjects. Sequences began at length 4 and increased to length 10, with two sequences presented at each length, for a total of 14 trials per run. Acquisition consisted of two runs, so that subjects reproduced a total of 28 different sequences generated by one of the two grammars.

Test Phase. Subjects proceeded seamlessly from the acquisition phase into the test phase, without being informed about the existence of the grammar. The task during testing was identical to the task during acquisition: Subjects reproduced test sequences by pressing the response buttons on the Simon memory game. For each subject, the same stimulus presentation format was used in both acquisition and testing. Half of the sequences used during testing were novel sequences that came from the grammar on which the subject had been trained. The other half of the testing sequences came from the grammar on which the subject had not been trained. "Trained" and "Not Trained" sequences were randomly distributed throughout the test phase. Sequences began at length 4 and increased to length 10, with two sequences presented at each length, for a total of 14 trials per run. The test phase consisted of four runs, so that subjects reproduced a total of 28 different sequences from the "Trained" grammar and 28 different sequences from the "Not Trained" grammar.

Recognition Test. After completing the acquisition and test phases, subjects were given instructions about making old-new and remember-know judgments. They were told that they would be given additional sequences on Simon, some of which they had been exposed to earlier during the first part of the experiment. If the sequence had been previously presented, subjects were to indicate "old" by pressing the Green button on Simon. If the sequence had not been previously presented, subjects were to indicate "new" by pressing the Red button on Simon. If a sequence was judged old, subjects were told to further distinguish whether they "remembered" the sequence or whether they "knew" that the sequence was old. Detailed instructions about making remember-know judgments were given, modeled after Rajaram (1993) and Roediger and McDermott (1995). Subjects were told that a "remember" judgment should be made for sequences for which they had a vivid memory of the actual original presentation of that sequence, while a "know" judgment should be made for sequences that they were confident had been presented, but for which they lacked the feeling of remembering the actual occurrence of the sequence. Subjects indicated a "remember" response by pressing the Yellow button on Simon and a "know" response by pressing the Blue button on Simon. To help subjects remember which button corresponded to which response, a card labeling each button with the appropriate response was placed on the Simon response box. Once again, subjects were not informed about the existence of the grammars prior to the recognition memory test.

The recognition test was composed of 50 sequences. Ten "Old Trained" and ten "Old Not Trained" sequences were selected from sequences used during the test phase. Ten "New Trained" and ten "New Not Trained" sequences were generated from each grammar, respectively, and presented as grammatical distracter items. Additionally, ten "New Random" sequences that did not conform to the rules of either grammar were used as distracters. Table 3 illustrates the five conditions used in the recognition memory test. The sequences used in the recognition memory test were of lengths 5 through 7, so that recognition performance would not be confounded by capacity limitations in memory span. Two sequences of length 5, four sequences at length 6, and four sequences at length 7 were used in each condition. The sequences in each condition were randomly distributed throughout the recognition test.

	Trained	Not Trained	Random
Old	10	10	
New	10	10	10

Table 3. Schematic of the distribution of sequences used in the recognition memory task.

After the recognition memory test was completed, subjects answered a set of questions attempting to assess their explicit knowledge of the grammars used to generate the sequences (modeled after Reber, 1989, 1993).

Results

Memory Span

Memory span scores were obtained by adding up the total number of items correctly reproduced on each perfectly recalled trial (an "absolute span score", after LaPointe & Engle, 1990). This scoring method was chosen because it provides a way of combining both list-based and item-based performance into a single composite score.

The mean memory span scores for the digit span task, the acquisition phase, and the test phase are shown in Table 4, listed by presentation modality group (n = 40 in each group). Two scores were obtained during the test phase, one score for sequences from the "Trained" grammar, and one score for sequences from the "Not Trained" grammar. Performance on the digit span task and the acquisition phase were comparable across all three modality groups. During the test phase, memory span scores for sequences that came from the "Trained" grammar were higher than memory span scores for sequences from the "Not Trained" grammar.

A one-way between-subjects ANOVA was performed on the digit span scores. The results revealed no differences among the three presentation modality groups ($F < 1$). The mean scores, ranging from 48.35 to 49.75, are consistent with the findings from other studies carried out in our laboratory that have used the same methods of measuring digit span (Goh & Pisoni, 1998; Karpicke & Pisoni, 2000). An additional one-way between-subjects ANOVA was performed on the acquisition phase scores. The results showed no differences among the three presentation modality groups ($F < 1$), indicating that performance during the acquisition phase was comparable in each presentation modality group.

Presentation Modality	Digit Span	Acquisition Phase	Test Phase: Trained	Test Phase: Not Trained
AV	48.55 (19.75)	68.98 (28.28)	83.65 (26.55)	63.45 (21.04)
AO	49.75 (19.39)	64.38 (25.73)	75.53 (28.28)	60.88 (26.99)
VO	48.35 (19.31)	63.40 (22.94)	70.90 (27.50)	60.65 (20.44)

Table 4. Mean memory span scores (standard deviation in parentheses) obtained from the digit span task, from the acquisition phase, and from the test phase. Performance during the test phase is divided into two span scores, one score for the "Trained" grammar, and one score for the "Not Trained" grammar. The scores are listed by presentation modality, $n = 40$ in each group.

"Sequence type during training" (Trained or Not Trained) was submitted to an ANOVA as a within-subjects factor, with presentation modality group (AV, AO, or VO) and grammar learned during training (A or B) as between-subjects factors. The analysis revealed a main effect of Sequence Type, $F(1, 114) = 73.002$, $p < .0001$, which will be referred to as the "learning effect". The learning effect refers to the finding that during the test phase, memory span for Trained sequences was significantly higher than memory span for Not Trained sequences. The learning effect in the auditory + visual (AV) condition is depicted in Figure 4, illustrating that performance on "Trained" sequences was better than performance on "Not Trained" sequences across almost all list lengths. Moreover, the learning effect did not interact with the grammar learned during training, $F(1, 114) = 2.384$, $p > .10$. This finding indicates that the learning effect was not simply a result of sequences from one grammar being easier to learn and remember than sequences from the other grammar. The interaction between the learning effect and presentation modality was marginal, $F(2, 114) = 2.677$, $p = .07$.

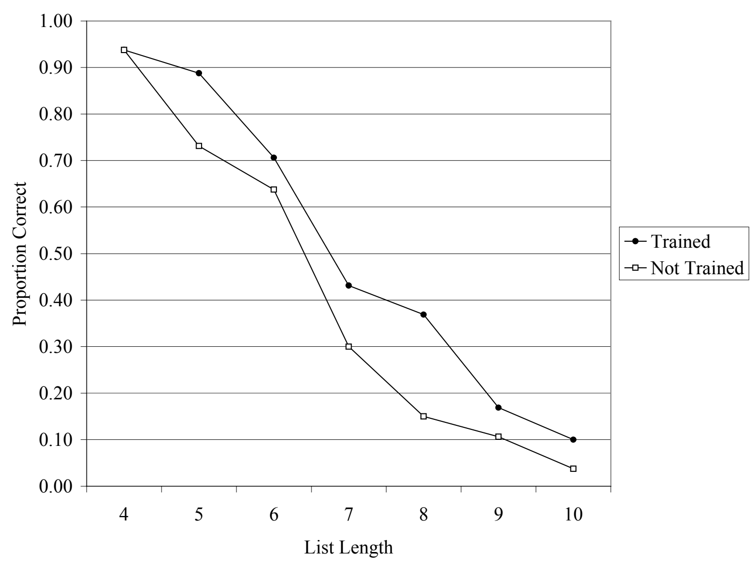


Figure 4. Depiction of the learning effect (memory span for Trained sequences is higher than memory span for Not Trained sequences) in the auditory+visual condition ($n = 40$), showing that performance on Trained sequences is better than performance on Not Trained sequences from list lengths 5 through 10.

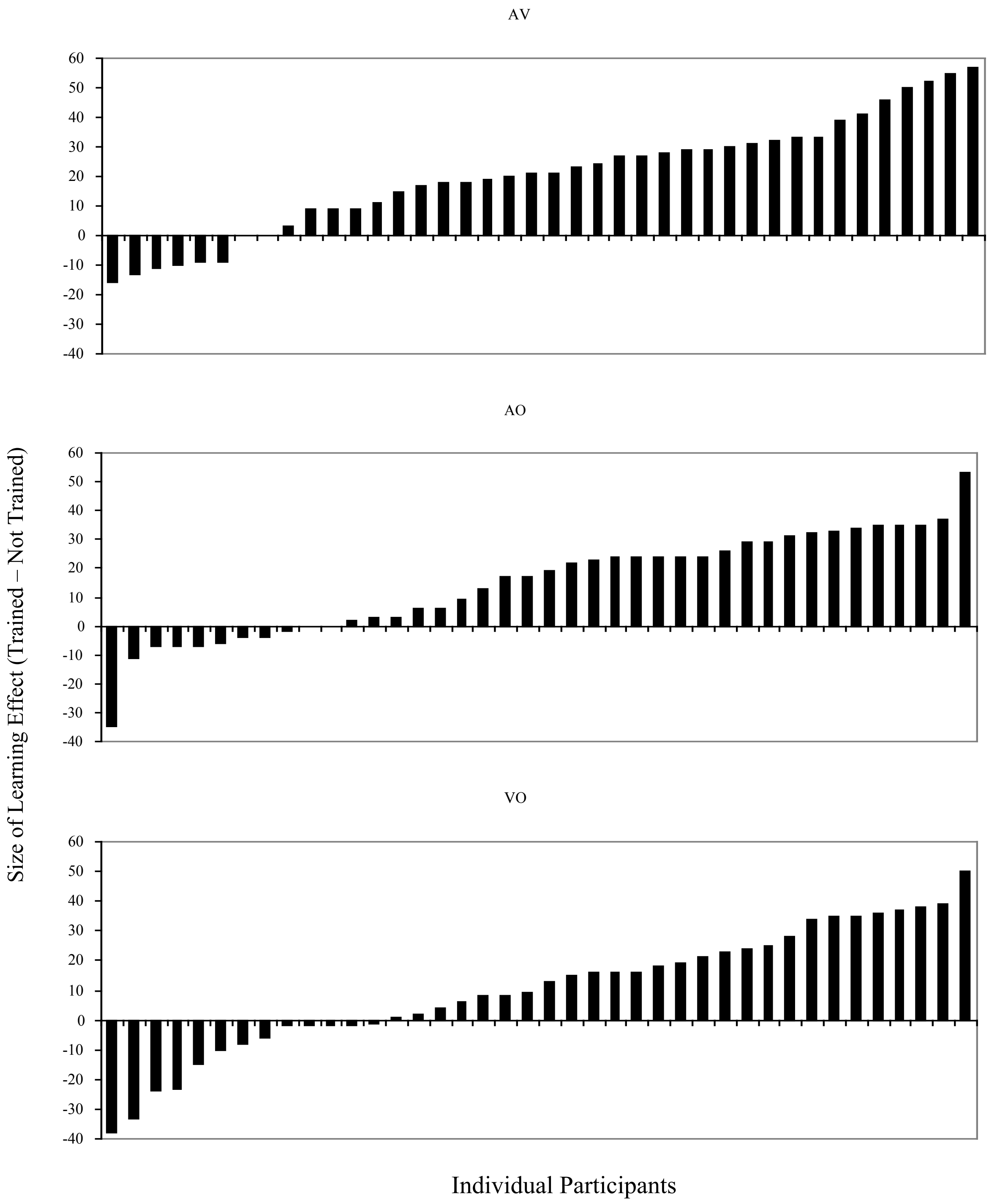


Figure 5. The size of the learning effect (Trained – Not Trained) for each individual participant in each of the three presentation modality conditions, rank ordered from smallest to largest effect. (AV=audiovisual, AO=auditory only, VO=visual only)

The size of the learning effect for each individual participant (i.e., "Trained" memory span score minus their "Not Trained" memory span score) is plotted in Figure 5. While most subjects showed a learning effect (88 out of 120, 73%), some did not, and there was a wide range of variability in the size of the learning effect among subjects across all three modality conditions (ranging from –38 to +57). To investigate some possible sources of this variability, bivariate correlations were performed comparing the size of the learning effect with digit span and with performance during the acquisition phase. Moderate positive correlations, shown in Table 5, were found between digit span and the learning effect, and between acquisition phase score and the learning effect, in both the auditory + visual (AV) and the auditory only (AO) conditions. In the visual only (VO) condition, a weak positive correlation was found between performance during the acquisition phase and the size of the learning effect ($p > .10$). No correlation was found between digit span and the learning effect in the VO condition.

Presentation Modality	Memory Span	Size of Learning Effect (r)
AV	Acquisition Phase	.52**
	Digit Span	.42**
AO	Acquisition Phase	.56**
	Digit Span	.44**
VO	Acquisition Phase	.20
	Digit Span	-.02

Table 5. Correlations of digit span and performance during the acquisition phase with the size of the learning effect. ** indicates $p < .01$.

The finding that performance during the acquisition phase correlated with the size of the learning effect is not entirely surprising, because it was predicted that individuals who performed better while acquiring the grammar would subsequently show a larger learning effect. However, the finding that an unrelated measure of memory span, the digit span task, correlated with the learning effect was more surprising. Figure 6 shows the scatter plots of the relationship between digit span and the size of the learning effect. These plots clearly illustrate that in the AV and AO conditions, individuals with higher auditory digit spans showed a larger learning effect.

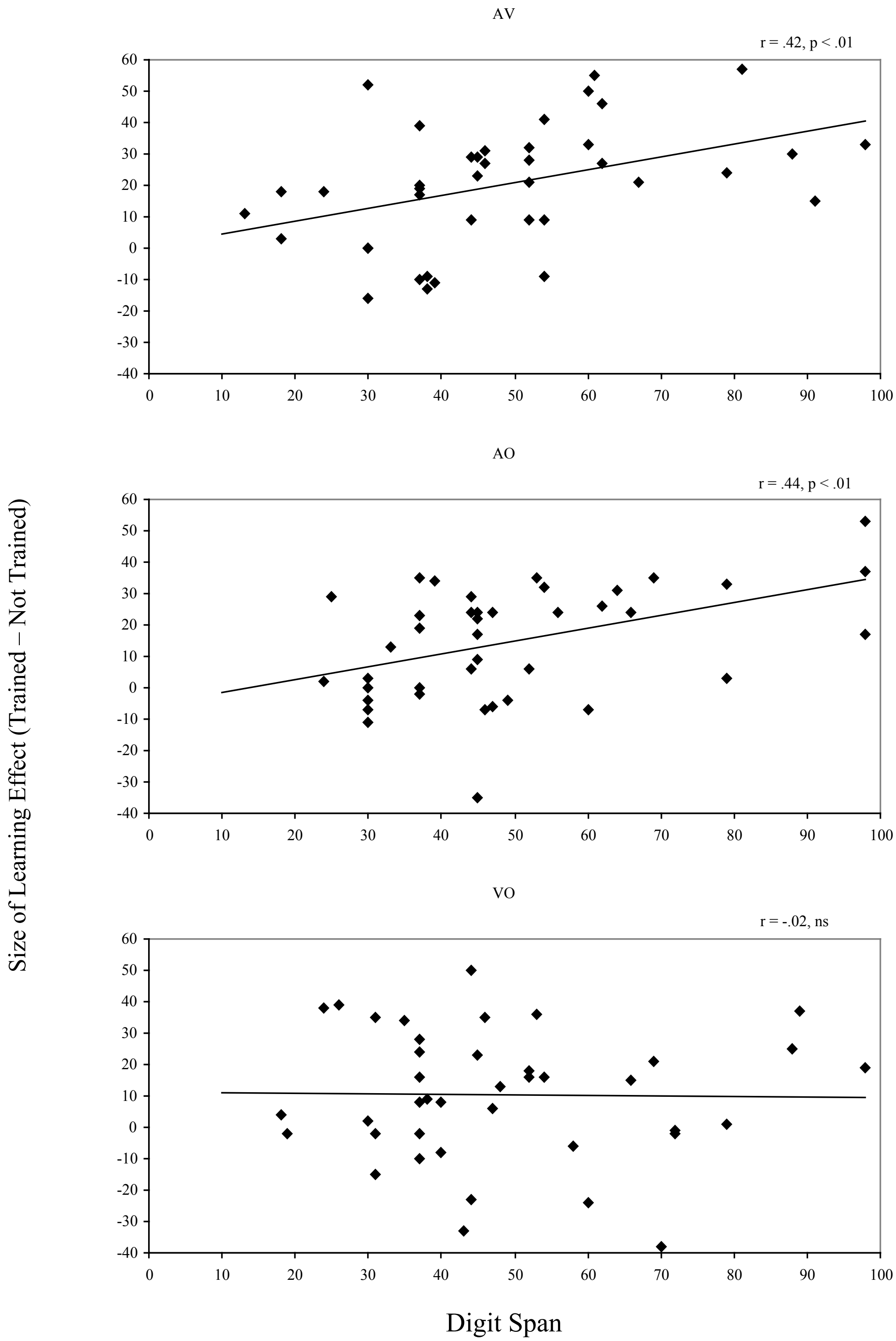


Figure 6. Scatter plots showing the relationship between auditory digit span and the size of the learning effect in each modality condition. (AV=audiovisual, AO=auditory only, VO=visual only)

Recognition Memory

Table 6 shows the probability of calling a sequence "Old" in the recognition memory test for each sequence type (Old Trained, Old Not Trained, New Trained, New Not Trained, and New Random; see Table 3). No differences were found in the pattern of hit rates and false alarm rates in all three presentation modality conditions. Thus, the data from the recognition memory test displayed in Table 6 are collapsed across modality groups. It is immediately evident that there were high false alarm rates for both "Trained" and "Not Trained" sequences, while subjects seemed to be able to correctly reject new random sequences with relative ease.

Sequence Type and Condition	Proportion of Responses Old	R	K
Old			
Trained	.54	.22	.32
Not Trained	.53	.20	.33
New			
Trained	.52	.23	.29
Not Trained	.47	.15	.32
Random	.28	.09	.19

Table 6. The probability of calling a sequence "Old" in the recognition memory test, collapsed across modality groups. R indicates "Remember" responses, K indicates "Know" responses.

To determine the relative discriminability of sequences from the "Trained" grammar and sequences from the "Not Trained" grammar, a measure of *d*' was obtained for the Trained and Not Trained conditions. Six subjects (out of 120 total) who performed either at floor or at ceiling in at least one of the recognition memory test conditions were excluded from the signal detection analysis. The mean *d*' for "Trained" sequences was 0.02, while the mean *d*' for "Not Trained" sequences was 0.18. Both of these numbers are very low, indicating an overall difficulty in discriminating old and new sequences generated by either grammar. Nevertheless, *d*' for "Not Trained" sequences was significantly higher than *d*' for "Trained" sequences, $t(113) = 2.414$, $p < .01$, indicating that subjects were worse at discriminating between old and new sequences generated by the grammar they were exposed to during acquisition.

Additional analyses of remember-know judgments made during the recognition memory test (see column two in Table 6) revealed a compelling false memory effect. The probability of saying "remember" to a sequence generated by the grammar on which the subject was "Not Trained" was significantly lower for new sequences than for old sequences, $t(119) = 2.873$, $p < .01$. However, the probability of saying "remember" to sequences generated by the "Trained" grammar was not different for old and new sequences, $t(119) = .396$, ns. This finding indicates that subjects falsely remembered sequences that had never been presented before, but which were generated by the grammar on which they had been trained, at about the same level as they correctly remembered sequences that had actually been presented previously.

Discussion

The primary findings from this study can be summarized as follows. First, a learning effect was found. Memory span for sequences generated by an artificial grammar to which subjects had been previously exposed was significantly higher than memory span for sequences generated by a different grammar that subjects had not been exposed to during the acquisition phase. This learning effect was found under three presentation conditions: auditory + visual (AV), auditory only (AO), and visual only (VO). Furthermore, the learning effect was not merely the result of one particular grammar generating sequences that were easier to memorize than those generated by the other grammar: Training on either grammar produced an equivalent subsequent learning effect in terms of increases in immediate memory capacity in the test phase.

Second, a wide range of variability in the size of the learning effect was found among individual subjects. Most subjects showed evidence of having benefited from prior exposure to one of the artificial grammars, but a few failed to show any benefit. Correlation analyses revealed that subjects who performed better during the initial acquisition phase, in which they were exposed to one of the artificial grammars, showed a larger learning effect during the test phase than subjects who performed more poorly. Furthermore, in the AV and AO conditions, significant correlations were found between performance on the digit span task, an unrelated measure of memory span capacity, and the size of the observed learning effect. This finding suggests that individuals with larger immediate memory span capacity are better able to acquire knowledge about the higher-order sequential dependencies underlying the structure of an artificial grammar and subsequently use their knowledge about the deep structure of the grammar to improve their memory span for new sequences that follow that particular set of grammatical constraints.

Finally, the recognition memory test revealed that subjects demonstrated learning effects without having the ability to display explicit knowledge about grammatical sequences. Although subjects were largely unable to discriminate between old and new sequences that had been generated by either grammar, subjects showed significantly better discrimination for sequences from the “Not Trained” grammar than for sequences from the “Trained” grammar. This finding indicates that implicitly learned knowledge about the grammar interfered with subjects’ ability to discriminate between old and new sequences generated by the grammar on which they had been trained. In addition, remember-know judgments obtained during the recognition memory test provided converging evidence about the subjects’ phenomenological experience of recollection. For sequences generated by the “Trained” grammar, the proportion of “remember” responses was essentially equivalent for old and new sequences. In other words, subjects claimed to have vivid memories for sequences that were generated by the system of grammatical rules on which they had been trained, but had never actually occurred anywhere previously during the course of the experiment.

The Process Dissociation Procedure Revisited

The primary motivation behind this study was to develop an experimental method of measuring implicit learning, based on the logic of the process dissociation procedure, that would be more sensitive to individual differences than standard implicit learning tasks. The critical assumption of the process dissociation procedure is that a given task is not a pure measure of a single underlying process (Jacoby, 1991). Instead of equating a task with a process, a better method for determining the relative contribution of distinct processes (e.g., automatic vs. intentional, unconscious vs. conscious, implicit vs. explicit) to performance on a given task is to design two paradigms, one in which the underlying processes work in concert to facilitate performance, and another in which the underlying processes work in opposition to interfere with performance.

In the present study, prior exposure to sequences generated by an artificial grammar facilitated subjects' ability to reproduce new sequences generated by the same grammar in an immediate memory span task. The pattern of results obtained suggests that implicit learning of the higher-order sequential dependencies that compose the artificial grammar was responsible for this facilitation in performance. If improved performance had been the result of some form of explicit knowledge about sequences generated by the artificial grammar, then such explicit knowledge should have also facilitated performance in an explicit recognition memory test. However, performance in the recognition memory test as indexed by *d*' was lowest for sequences generated by the grammar to which subjects had been initially exposed, suggesting that implicitly learned knowledge about the underlying structural regularities of the grammar acted in opposition with explicit knowledge to interfere with recognition performance.

The purpose of the present study was not to invent an "implicit learning task." In fact, the argument presented here is that there is no such thing as an implicit learning task. Instead, this study has explored one possible method of measuring the implicit influences on a largely explicit task, involving immediate memory span. The demonstration that two knowledge sources, one implicit and the other explicit, can act either in concert to facilitate performance or in opposition to interfere with performance provides strong support for a process dissociation without relying on a task dissociation. Implicitly learned knowledge selectively improved memory span performance (E+I) and at the same time selectively interfered with recognition memory performance (E–I). Using this procedure, the relative contribution of implicitly learned knowledge to memory span performance was determined for each individual subject and used to explore variation in performance on these tasks. This index of implicit learning revealed a wide range of individual differences not found in previous investigations of implicit learning, which have equated performance on the standard artificial grammar learning task with the degree of implicit learning (Mayberry et al., 1995; McGeorge et al., 1997; Reber et al., 1991). The present investigation differs from these earlier studies in that the primary measure of implicit learning was based on peak immediate memory span, where large individual differences are typically observed.

Relating Capacity and Sequential Learning

The present findings provide a serious challenge to any theory of human cognition that simply writes off individual differences in implicit cognitive processes as non-existent (e.g., Reber, 1993, 1997a). This study has provided new empirical evidence that implicit learning can vary greatly among individuals. When implicit learning is measured not by performance on a given task but, rather, as a process that makes some relative contribution (either positive or negative) to task performance, the variability in implicit learning among individuals is revealed. In addition to this empirical evidence, there are sound theoretical reasons for arguing that individual differences in implicit learning might be responsible in part for individual differences in other cognitive abilities as well. For example, if the mechanism of implicit learning is a substrate for language abilities, then one would expect to find substantial co-variation among a large range of language measures as well as among measures of implicit learning.

The nature of individual differences in implicit learning is not well understood at this time and will require more extensive investigation. The findings in the present study, however, suggest that a relationship exists between memory span capacity and the ability to learn and to use higher-order sequential dependencies. At least one model, proposed by Elman (1993), has also suggested a relationship between memory capacity and the mechanisms of sequential learning. In his study, Elman (1993) constructed a "semi-realistic" artificial language and then attempted to train a simple recurrent network to process sentences from this language. A simple recurrent network is a special type of artificial neural network that employs a "context layer," which is essentially a feedback mechanism, to provide the network with additional input about its own prior internal states (see also Elman, 1990). This recurrent property of the network renders it capable of processing sequential input and learning sequential dependencies by allowing it to encode information about context. In Elman's model, memory capacity was represented by the access the network had to its own prior internal states via the recurrent connections. In order to simulate the maturational development of memory span during childhood, Elman created a network that began with a limited memory capacity that gradually increased by having greater access to the recurrent connections. Elman found that these conditions were optimal for the neural network to learn the "long-distance dependencies" of the artificial language. Elman argued that there is a direct relationship between the gradual increase in memory capacity during childhood and the ability to learn and process sequential information.

The suggestion made in Elman's (1993) simple recurrent network model is that a system with a small memory capacity first learns "shorter-distance" dependencies in a language. As capacity increases, the network becomes able to process "longer-distance" (i.e., "higher-order") sequential dependencies by building upon its prior knowledge about shorter-distance dependencies. Based on these results, Elman (1993) proposed a direct relationship between memory capacity and sequential learning, that the ability to learn higher-order sequential dependencies is a function of the size of memory capacity.

Future Directions

Future studies using immediate memory span to measure implicit learning should be able to address a wide variety of issues. To investigate the representational form of implicitly learned knowledge, a cross-modal transfer study could be carried out using the process dissociation procedure. As mentioned earlier, Altmann et al. (1995) and Manza and Reber (1997) have shown that implicit knowledge acquired in an artificial grammar learning task can be transferred across sensory modalities. An extension of their findings would be to investigate whether exposure to a grammar in one modality (auditory or visual) would transfer over and improve immediate memory span for grammatical sequences presented to the other modality. Finding that the acquisition of implicit knowledge in one modality can be applied to improve memory span for information in another modality would provide further support for abstractionist views of cognition which argue that implicitly learned knowledge is not bound to the physical form of the stimulus but is a representation of higher-order dependencies.

While we showed that individual differences in implicit learning are related to memory span capacity, we did not investigate any possible relationships between implicit learning and individual differences in working memory capacity or executive attention (Engle, 2001, 2002). Working memory capacity is thought to be a system for keeping goal-relevant information active in memory. This construct has been shown to be functionally distinct from traditional views of short-term memory, in that working memory capacity is essentially short-term memory plus controlled attention (Engle, Tuholski, Laughlin, & Conway, 1999; Kail & Hall, 2001). Measures of working memory capacity, such as the operation span, are thought to measure a different cognitive ability than measures of short-term memory capacity, such as the digit span. An investigation of the relationship between working memory and implicit learning using measures of immediate memory span could better clarify the role of attention, the ability to maintain goal-relevant information in an active state, in learning sequential dependencies.

Another important direction for future work would be to investigate relationships between implicit learning and individual differences in language abilities. If the mechanisms of implicit learning are fundamental for language learning, then direct comparisons between implicit learning and any measures of language abilities would clarify this relationship. Among both adults and children, there is a wide range of variability in measures of language processing. Such measures include word recognition, vocabulary knowledge, reading, language comprehension, speaking rate, speech production, second language learning, and "metalinguistic" skills (i.e., what you know that you know about a language). Though implicit learning may be responsible for the observed individual differences in several measures of language processing abilities, no research program has focused specifically on exploring these possible relationships.

Language is an example of a behavior that is retained over a long period of time. Few studies in the implicit learning literature have investigated the long-term retention of implicit knowledge. In one study, Allen and Reber (1980) recruited subjects who had participated in an artificial grammar learning study two years prior (Reber & Allen, 1978) to assess what they had retained from their earlier grammar learning experience. Allen and Reber (1980) found that even two years after learning an artificial grammar, subjects could still correctly classify grammatical and non-grammatical letter strings without having to "re-learn" the grammar. A similar experiment, conducted using the process dissociation procedure, could investigate how well implicit learning is retained over time, whether individual differences in implicit learning can gradually change, and perhaps even whether there is some means of improving an individual's ability to implicitly acquire knowledge about a complex environment.

One of the predictions derived from Reber's (1992a,b, 1993) model of the cognitive unconscious is that implicit learning should be relatively age-independent, compared with explicit cognitive processes. However, this prediction was made based on empirical evidence from the body of knowledge about implicit learning at the time. The process dissociation procedure applied in the present experiment, using immediate memory span to measure implicit learning, proved to be more sensitive to individual differences than the methodology used in standard artificial grammar learning studies. Thus, using immediate memory span as a measure of implicit learning might also be more sensitive to variation across different age groups and different developmental levels. Individual differences in implicit learning might be related to some of the language difficulties experienced by children, such as language delays, for example. In addition, individual differences in implicit learning might account for some portion of the well-documented age-related declines found in a wide range of various cognitive abilities (e.g., Salthouse, 1996).

Finally, the present study has some direct and immediate implications concerning an important clinical problem currently being investigated in our laboratory on the nature of individual differences in deaf children with cochlear implants. As mentioned earlier, there is a wide range of individual differences in outcome on measures of language abilities in this population of children, without any clear explanation of why some children do very well at acquiring language with their implants while others do not (Pisoni & Cleary, 2002). Research in our laboratory has investigated whether individual differences in memory span capacity can account for some portion of the variance in such language measures (Cleary, Pisoni, & Geers, 2001). The present study provides evidence for a direct relationship between memory span capacity and implicit learning, and provides converging evidence supporting the proposal that implicit learning might be a critical mechanism for language learning (see also Winter & Reber, 1994). The new method of measuring implicit learning developed in this study proved to be more sensitive to individual differences than other traditional measures, such as the standard artificial grammar learning task. Perhaps differences in fundamental abilities, such as the ability to learn and to exploit higher-order sequential dependencies in any of various complex stimulus environments, can help explain why some children acquire language very well with their cochlear implants while others do not. Investigation of the possible relationships between implicit learning and individual differences in language abilities in deaf children with cochlear implants could have a wide range of implications, from possibly explaining some of the observed variability in language development, to potentially predicting an individual child's outcome before implantation, and to assessing the benefits of new intervention methods.

References

Allen, R., & Reber, A.S. (1980). Very long term memory for tacit knowledge. *Cognition, 8*, 175-185.

Altmann, G.T.M., Dienes, Z., & Goode, A. (1995). Modality-independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21,* 899-912.

Brody, N. (1989). Unconscious learning of rules: Comment on Reber's analysis of implicit learning. *Journal of Experimental Psychology: General, 118*, 236-238.

Brooks, L.R., & Vokey, J.R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*, 316-323.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: Wiley.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, N., & Miller, G.A. (1958). Finite state languages. *Information and Control, 1,* 91-112.

Cleary, M., Pisoni, D.B., & Geers, A.E. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear & Hearing, 22*, 395-411.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8*, 343-350.

Dulany, D.E., Carlson, R.A., & Dewey, G.I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General, 113*, 541-555.

Dulany, D.E., Carlson, R.A., & Dewey, G.I. (1985). On consciousness in syntactical learning and judgment: A reply to Reber, Allen, and Regan. *Journal of Experimental Psychology: General, 114*, 25-32.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48,* 71-99.

Engle, R.W. (2001). What is working memory capacity? In H.L. Roediger, J.S. Nairne, I. Neath, & A.M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297-314). Washington D.C.: American Psychological Association.

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

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Fodor, J.A. (1983). *Modularity of mind*. Cambridge: MIT Press.

Goh, W.D., & Pisoni, D.B. (1998). Effects of lexical neighborhoods on immediate memory span for spoken words: A first report. *Research on Spoken Language Processing Progress Report No. 22* (pp. 195-213). Bloomington, IN: Speech Research Laboratory, Indiana University.

Graf, P., & Schacter, D.L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 501-518.

Gupta, P., & Cohen, N.J. (in press). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review*.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Hsaio, A.T., & Reber, A.S. (1998). The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious. In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of implicit learning* (pp. 201-222). Thousand Oaks, CA: Sage Publications, Inc.

Jacoby, L.L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22,* 485-508.

Jacoby, L.L. (1988). Memory observed and memory unobserved. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 145-177). New York: Cambridge University Press.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513-541

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306-340.

Jacoby, L.L., Toth, J.P., & Yonelinas, A.P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*, 139-154.

Kail, R., & Hall, L.K. (2001). Distinguishing short-term memory from working memory. *Memory & Cognition, 29*, 1-9.

Karpicke, J., & Pisoni, D.B. (2000). Memory span and sequence learning using multimodal stimulus patterns: Preliminary findings in normal-hearing adults. *Research on Spoken Language Processing Progress Report No. 24,* (pp. 393-406). Bloomington, IN: Speech Research Laboratory, Indiana University.

Kihlstrom, J.F. (1987). The cognitive unconscious. *Science, 237*, 1445-1452.

Kirsner, K. (1994). Second language vocabulary learning: The role of implicit processes. In N.C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 283-311). London: Academic Press.

Knopman, D.S., & Nissen, M.J. (1987). Implicit learning in patients with probable Alzheimer's disease. *Neurology, 37*, 784-788.

Knowlton, B.J., & Squire, L.R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 79-91.

Knowlton, B.J., & Squire, L.R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 169-181.

LaPointe, L.B., & Engle, R.W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16,* 1118-1133.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252-271.

Manza, L., & Reber, A.S. (1997). Representing artificial grammars: Transfer across stimulus forms and modalities. In D.C. Berry (Ed), *How Implicit is Implicit Learning?* (pp. 73-106). Oxford: Oxford University Press.

Mathews, R.C., Buss, R.R., Stanley, W.B., Blanchard-Fields, F., Cho, J.R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1083-1100.

Mayberry, M., Taylor, M., & O'Brien-Malone, A. (1995). Implicit learning: Sensitive to age but not IQ. *Australian Journal of Psychology, 47*, 8-17.

McGeorge, P., Crawford, J.R., & Kelly, S.W. (1997). The relationships between psychometric intelligence and learning in an explicit and an implicit task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 239-245.

Miller, G.A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Miller, G.A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology, 56,* 485-491.

Miller, G.A. (1967). Project grammarama. In G.A. Miller, *The Psychology of Communication: Seven Essays.* New York: Basic Books.

Miller, G.A., & Chomsky, N. (1963). Finitary models of language users. In R.D. Luce, R.R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology* (pp. 419-491). New York: Wiley.

Miyamoto, R.T. (1995). Cochlear implants. *Implants in Otolaryngology, 28*, 287-294.
Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1-32.
Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 119*, 264-275.
Perruchet, P., & Pacteau, C. (1991). Implicit acquisition of abstract knowledge about artificial grammars: Some methodological and conceptual issues. *Journal of Experimental Psychology: General, 120,* 112-116.
Pinker, S. (1994). *The language instinct*. New York: William Morrow and Co.
Pisoni, D.B. (1999). Individual differences in effectiveness of cochlear implants in prelingually deaf children: Some new process measures of performance. *Research on Spoken Language Processing Progress Report No. 23* (pp. 3-49). Bloomington, IN: Speech Research Laboratory, Indiana University.
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. (2002). 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.), *Handbook of auditory research: Auditory Prostheses, SHAR Volume X*. New York: Springer-Verlag.
Pisoni, D.B., & Geers, A.E. (2000). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language. *Annals of Otology, Rhinology and Laryngology Supplement, 185*, 92-93.
Pisoni, D.B., Cleary, M., Geers, A.E., & Tobey, E.A. (2000). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *The Volta Review, 101*, 111-164.
Pisoni, D.B., Svirsky, M.A., Kirk, K.I., & Miyamoto, R.T. (1997). Looking at the "Stars": A first report on the intercorrelations among measures of speech perception, intelligibility, and language development in pediatric cochlear implant users. *Research on Spoken Language Processing Progress Report No. 21* (pp. 336-343). Bloomington, IN: Speech Research Laboratory.
Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*, 89-102.
Rajaram, S. & Roediger, H.L. (1997). Remembering and knowing as states of consciousness during retrieval. In J.D. Cohen & J.W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 213-240). Hillsdale: Erlbaum.
Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 855-863.
Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81*, 115-119.
Reber, A.S. (1976). Implicit learning of artificial grammars: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 88-94.
Reber, A.S. (1989). Implicit learning and tacit knowledge*. Journal of Experimental Psychology: General, 118*, 219-235.
Reber, A.S. (1990). On the primacy of the implicit: Comment on Perruchet and Pacteau. *Journal of Experimental Psychology: General, 119,* 340-342.
Reber, A.S. (1992a). The cognitive unconscious: An evolutionary perspective. *Consciousness & Cognition, 1*, 93-133.
Reber, A.S. (1992b). An evolutionary context for the cognitive unconscious. *Philosophical Psychology, 5*, 33-51.
Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Oxford University Press.

Reber, A.S. (1997a). How to differentiate implicit and explicit modes of acquisition. In J.D. Cohen & J.W. Schooler (Eds), *Scientific approaches to consciousness* (pp. 137-159). Hillsdale: Erlbaum.

Reber, A.S. (1997b). Caterpillars and consciousness. *Philosophical Psychology, 10*, 437-449.

Reber, A.S., & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition, 6*, 189-221.

Reber, A.S., & Allen, R. (2000). Individual differences in implicit learning: Implications for the evolution of consciousness. In R.G. Kunzendorf & B. Wallace (Eds.), *Individual differences in conscious experience* (pp. 227-247). Amsterdam: John Benjamins Publishing Company.

Reber, A.S., Kassin, S.M., Lewis, S., & Cantor, G.W. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 492-502.

Reber, A.S., Walkenfeld, F.F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 888-896.

Roediger, H.L. (1990a). Implicit memory: A commentary. *Bulletin of the Psychonomic Society, 28*, 373-380.

Roediger, H.L. (1990b). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043-1056.

Roediger, H.L., & McDermott, K.B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology*. Vol. 8 (pp. 63-131). Amsterdam: Elsevier.

Roediger, H.L., & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Roediger, H.L., Weldon, M.S., Stadler, M.A., & Riegler, G.H. (1992). Direct comparison of word stems and word fragments in implicit and explicit retention tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1251-1269.

Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403-428.

Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501-518.

Shanks, D.R., & St. John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367-448.

Stadler, M.A., & Roediger, H.L. (1998). The question of awareness in research on implicit learning. In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of Implicit Learning* (pp. 105-132). Thousand Oaks, CA: Sage Publications.

Squire, L.R. (1987). *Memory and brain*. New York: Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1-12.

Tulving, E., & Schacter, D.L. (1990). Priming and human memory systems. *Science, 247*, 301-306.

Vokey, J.R., & Brooks, L.R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 328-344.

Warrington, E.K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature, 217*, 972-974.

Warrington, E.K., & Weiskrantz, L. (1970). Amnesic syndrome: Consolidation or retrieval? *Nature, 228,* 628-630.

Willingham, D.B., Nissen, M.J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1047-1060.

Winter, B., & Reber, A.S. (1994). Implicit learning and natural language acquisition. In N.C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 115-146). London: Academic Press.

Woodworth, R.S. (1938). *Experimental psychology*. New York: Henry Holt Co.

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.

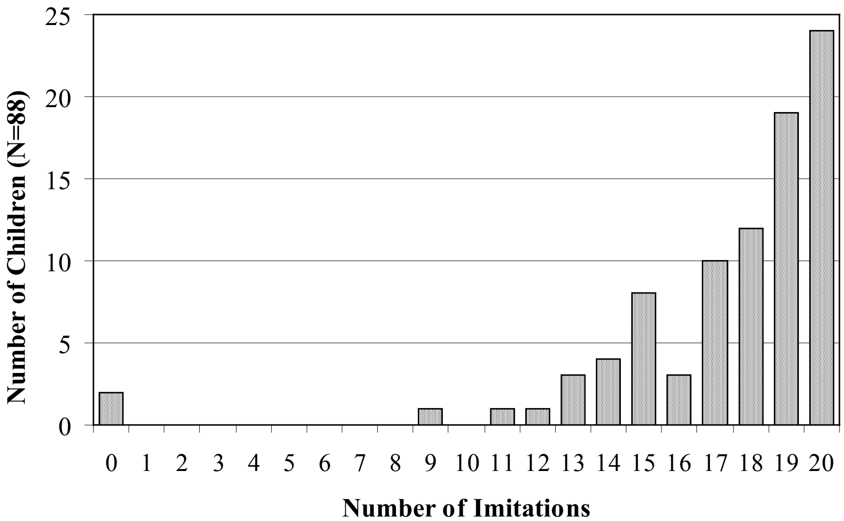


Figure 1. The number of imitations produced by the original 88 children who participated in the nonword repetition task during the 1999 and 2000 CID summer programs.

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
	ballop	ˈbæ.ləp
	prindle	ˈprɪn.dl̩
	rubid	ˈruˌbɪd
2	sladding	ˈslæ.diŋ
	tafflist	ˈtæ.flɪst
	bannifer	ˈbæ.nəˌfɚ
	berrizen	ˈbɛ.ɹəˌzɪn
	doppolate	ˈdɑ.pəˌle͡ɪt
3	glistering	ˈglɪ.stɚ.iŋ
	skiticult	ˈskɪ.ɾəˌkʌlt
	comisitate	kəˈmi.səˌte͡ɪt
	contramponist	kənˈtɹæm.pəˌnɪst
	emplifervent	ɛmˈplɪ.fɚˌvɛnt
4	fennerizer	ˈfɛ.nɚˌa͡ɪ.zɚ
	penneriful	pəˈnɛ.ɹəˌfʌl
	altupatory	ælˈtu.pəˌtɔ.ri
	detratapillic	diˈtræ.ɾəˌpɪ.lɪk
5	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) Segment Score: 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) Manner Feature Score: 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) Voice Feature Score: 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%).

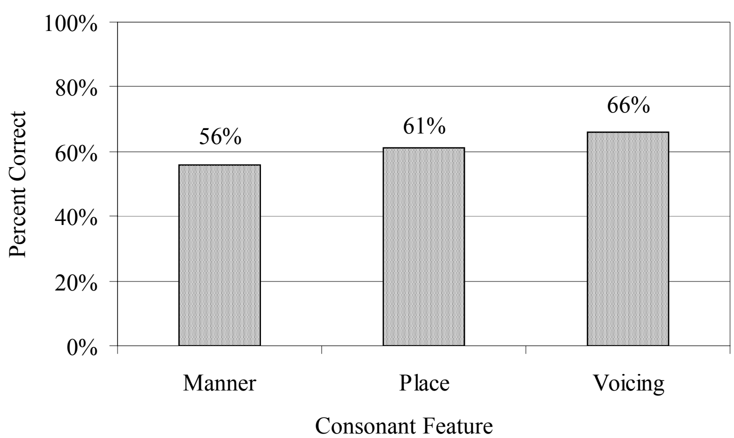


Figure 2. The percent of target consonants that were imitated correctly overall, and in terms of place, in terms of manner, and in terms of voicing.

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.

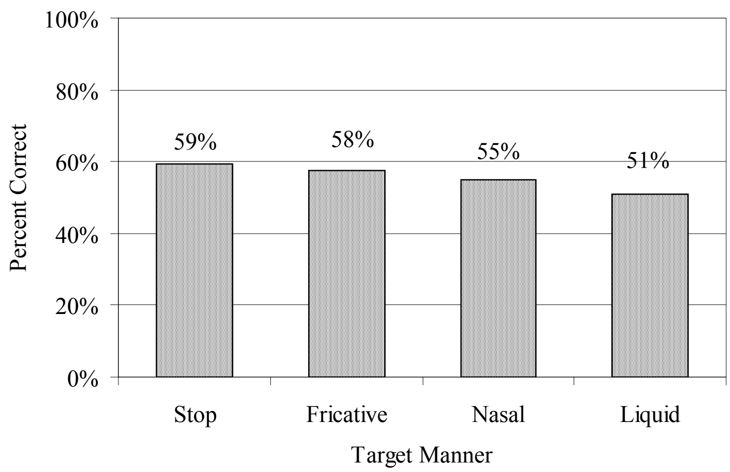


Figure 3. Percent of target consonants imitated correctly in terms of manner.

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.

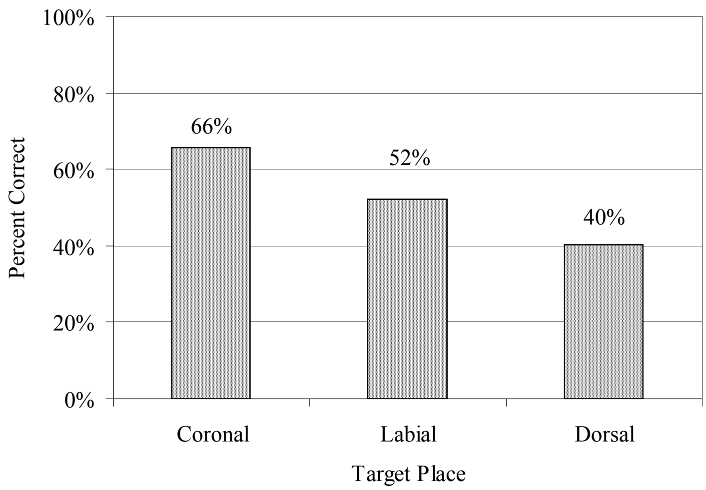


Figure 4. Percent of target consonants imitated correctly in terms of place.

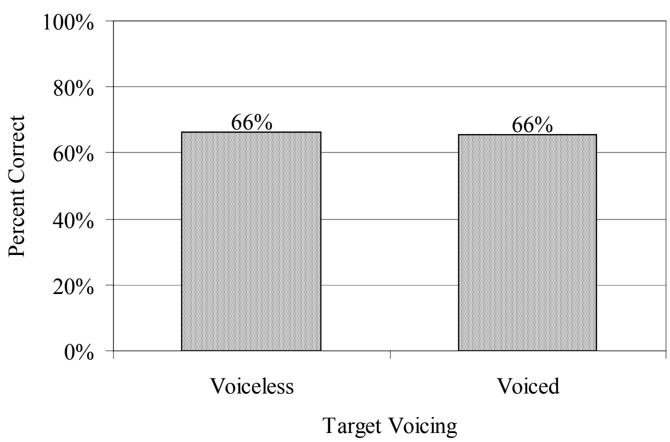


Figure 5. Percent of target consonants imitated correctly in terms of voicing.

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 ($p < .05$).

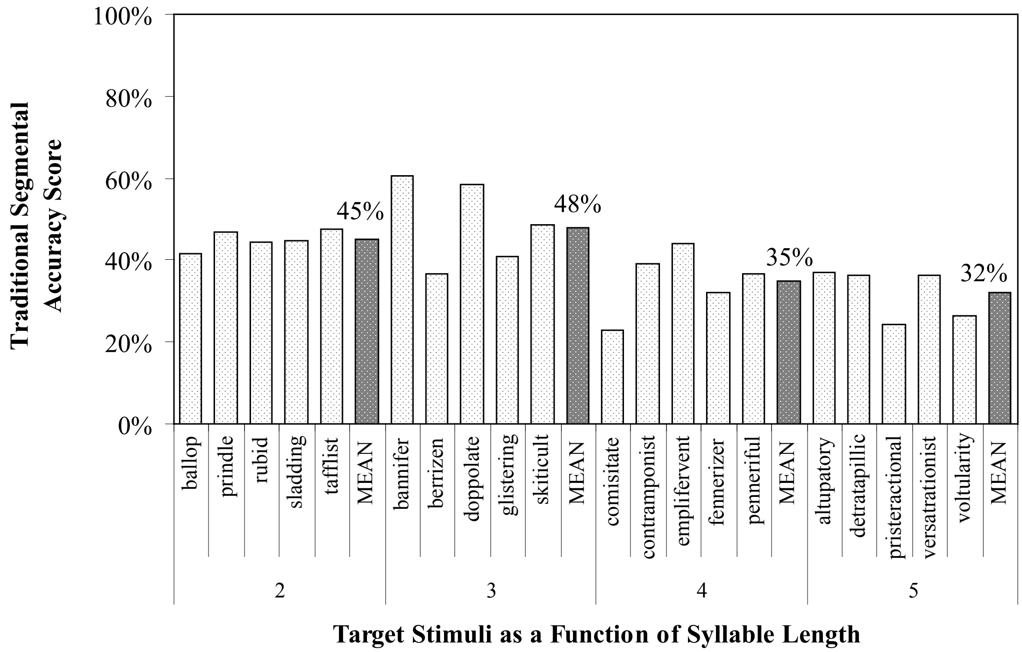


Figure 6. Mean traditional segmental accuracy score for each target nonword and for each target syllable length, averaged across children.

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%.

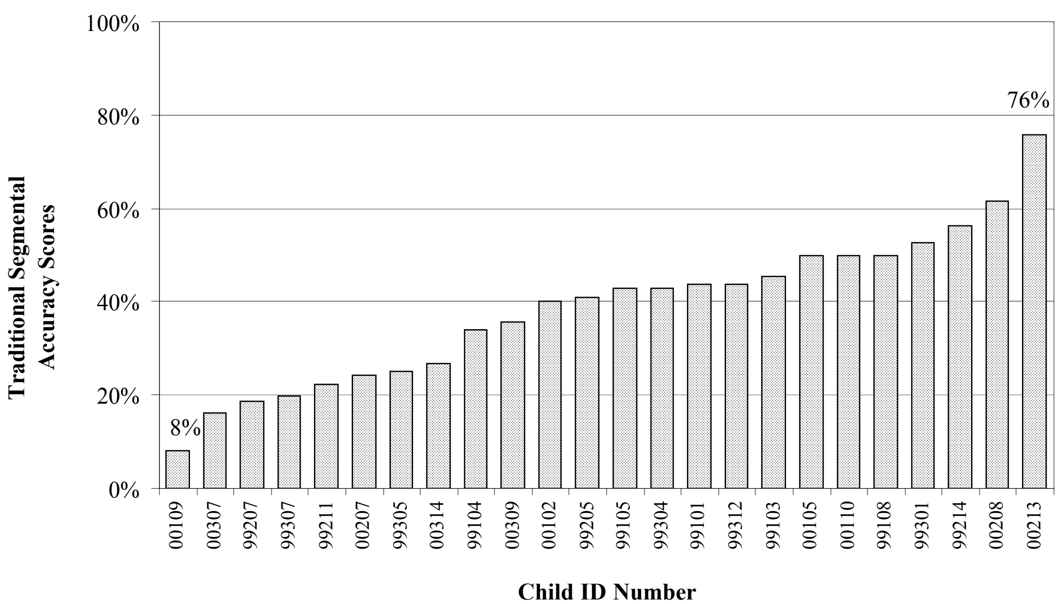


Figure 7. Mean traditional segmental accuracy score for each child, averaged across target nonwords.

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.

Outcome and Process Measures of Performance	Correlation *r*-values: 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.

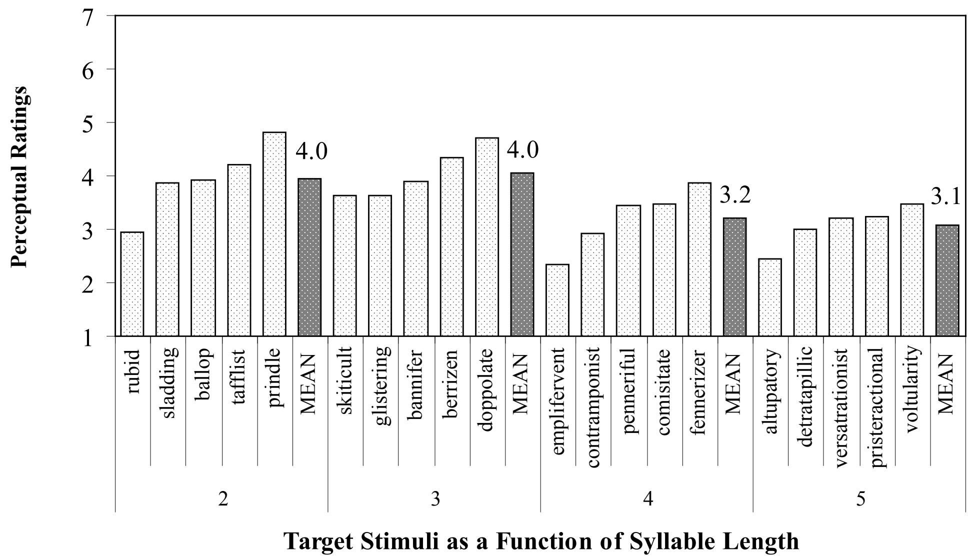


Figure 8. Mean perceptual rating for each target nonword and for each target syllable length, averaged across children.

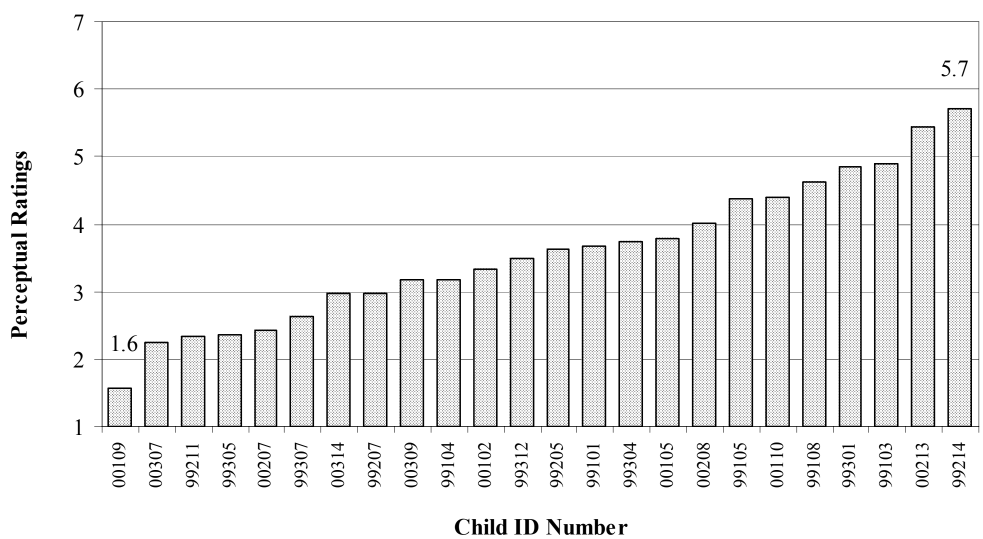


Figure 9. Mean perceptual rating for each child, averaged across target nonwords.

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.

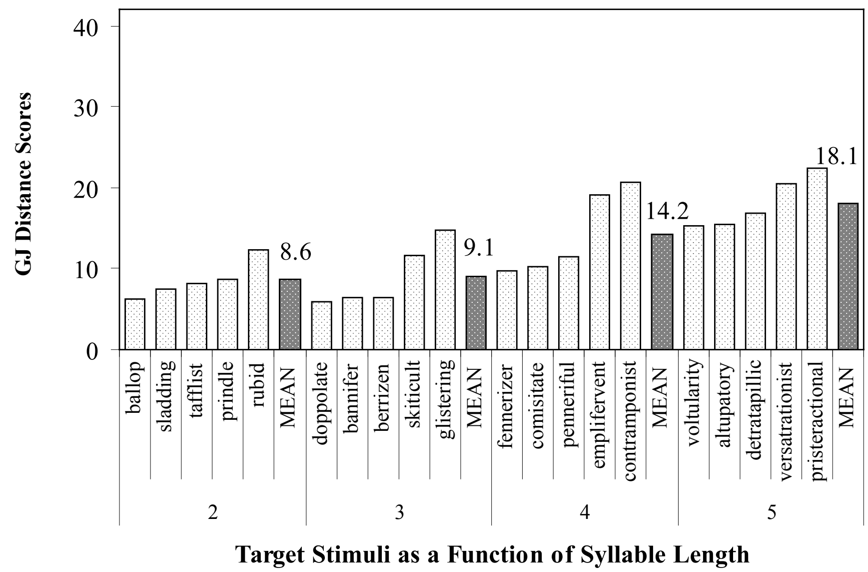


Figure 10. GJ distance score for each target nonword and for each target syllable length, averaged across children.

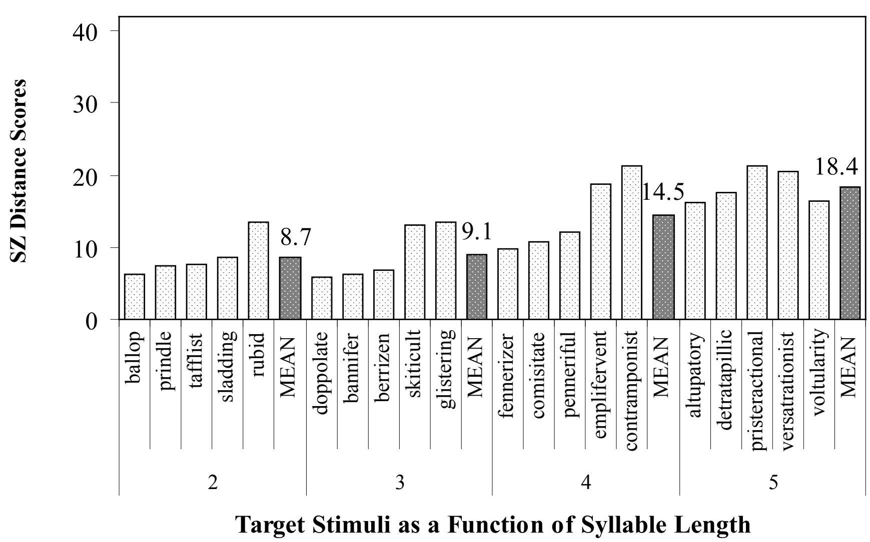


Figure 11. SZ distance score for each target nonword and for each target syllable length, averaged across children.

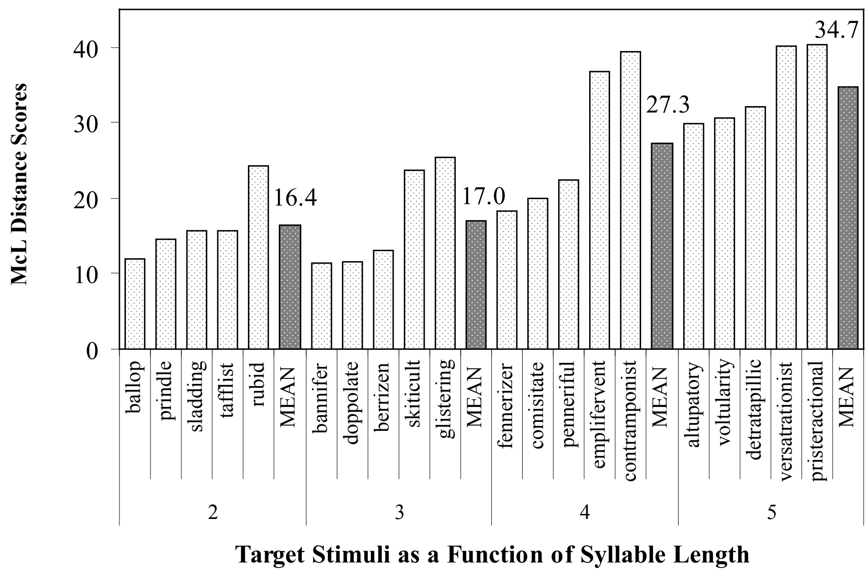


Figure 12. McL distance score for each target nonword and for each target syllable length, averaged across children.

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.

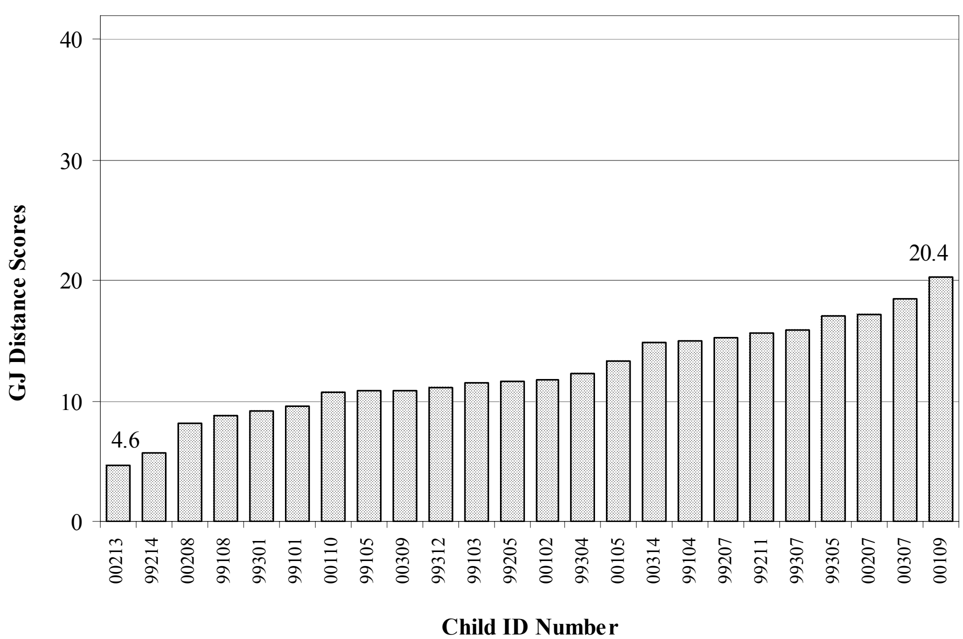


Figure 13. GJ distance score for each child, averaged across target nonwords.

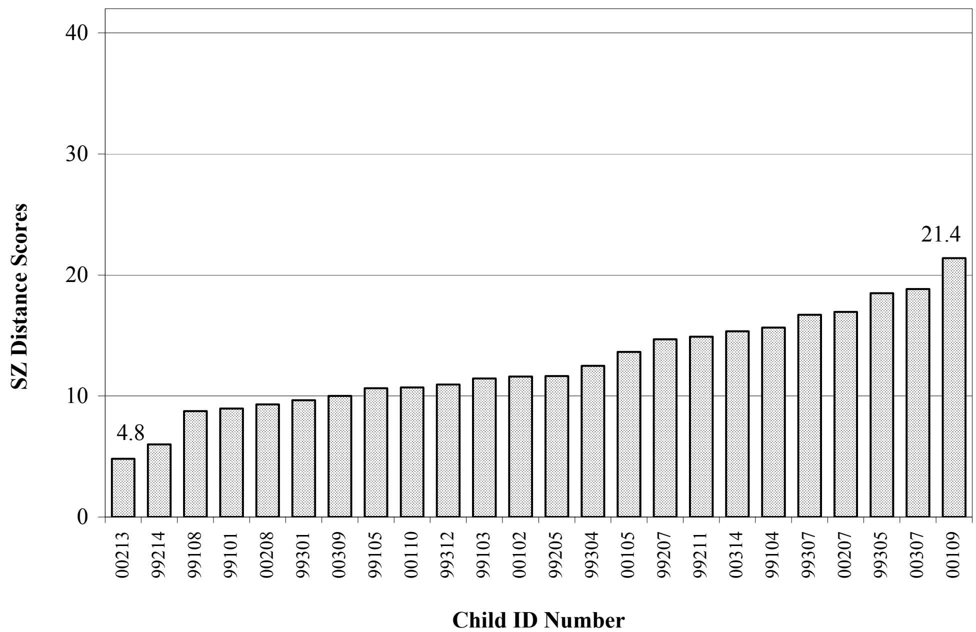


Figure 14. SZ distance score for each child, averaged across target nonwords.

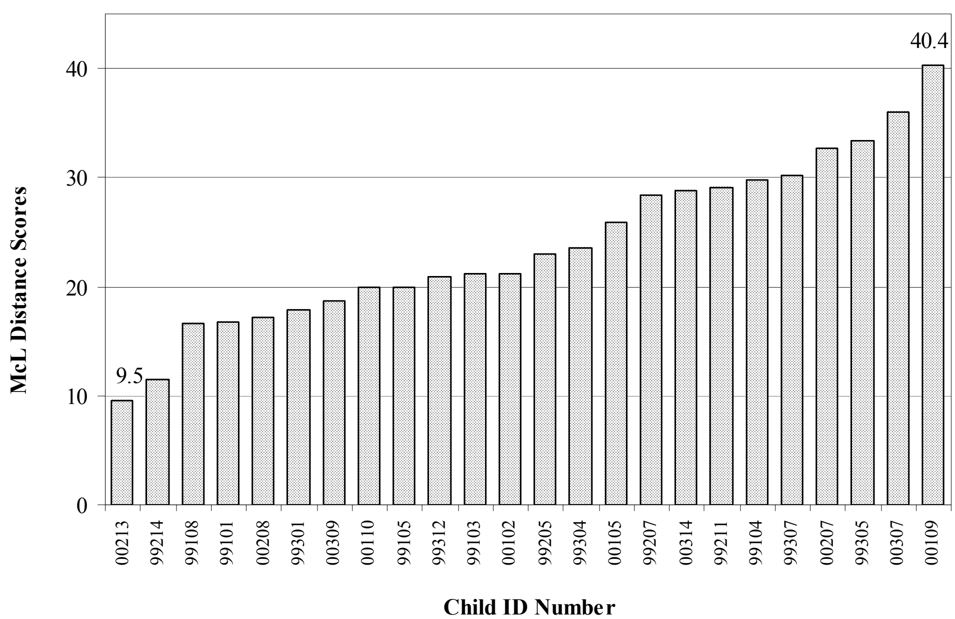


Figure 15. McL distance score for each child, averaged across target nonwords.

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 communicatio	- .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.

Outcome and Process Measures of Performance	Correlation *r*-values 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.

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

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

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

[1] This research was supported by NIH research grant DC00111 and NIH T32 training grant DC00012 from the NIDCD to Indiana University, Bloomington. We are grateful to Dr. Ann Geers, Rosalie Uchanski, Chris Brenner and the research team at the Center for Applied Research on Childhood Deafness, Central Institute for the Deaf, for their invaluable help on this project. We also wish to acknowledge the help and cooperation of Dr. Emily Tobey and her colleagues at Callier Advanced Hearing Research Center at the University of Texas, Dallas. Finally, we would also like to thank Miranda Cleary, Cynthia Clopper, and Rose Burkholder for their help with data collection and analysis, and other colleagues at Indiana University for useful comments and discussion.

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

Abstract. In this study, we examined two prosodic characteristics of speech production in eight- to 10-year-old experienced cochlear implant (CI) users who completed a nonword repetition task. We looked at how often they correctly reproduced syllable number and primary stress location in their responses. Although only 5% of all imitations were produced correctly without errors, 64% of the imitations contained the correct syllable number and 61% had the correct placement of primary stress. Moreover, these target prosodic properties were correctly preserved significantly more often for targets with fewer syllables and targets with primary stress on the initial syllable. Syllable and stress scores were significantly correlated with measures of speech perception, intelligibility, perceived accuracy, and working memory. These findings suggest that pediatric CI users encode the overall prosodic envelope of nonword patterns, despite the loss of more detailed segmental properties. This phonological knowledge is also reflected in other language and memory skills.

Introduction

Research on the speech production skills of pediatric cochlear implant (CI) users has shown that while they perform more poorly than normal-hearing children on a wide range of tasks, their speech and language performance improves significantly after receiving a cochlear implant. Previous studies have largely centered on segmental and featural aspects of children's phonological systems. For example, in terms of segments, deaf children with CIs tend to produce vowels more accurately than consonants (Tye-Murray, Spencer, Bedia, & Woodworth, 1996; Brown & McDowall, 1999; Serry & Blamey, 1999). In terms of features, it has generally been found that stops are produced more accurately than fricatives (Tobey, Geers, & Brenner, 1994; Serry & Blamey, 1999) and labials are produced more accurately than non-labials (Tobey et al., 1994; Serry, Blamey, & Grogan, 1997). Reports regarding the acquisition and accuracy of voicing are not as consistent. In one study, Tobey et al. found that during their first three years of implantation, children with CIs improved in their production of voiceless fricatives and both voiced and voiceless stops to a greater degree than voiced fricatives. More recently, Serry and Blamey reported that pediatric CI users preferred voiced consonants (i.e. they emerged earlier and the children acquired more of them) to voiceless consonants. It is possible that production of voicing, more than manner or place, is more dependent on other factors such as the surrounding segmental context or position within the syllable (Dillon & Cleary, 2000).

Studies of suprasegmental aspects of speech production are fewer in number and primarily concern phonetic properties of speech. For example, Tobey et al. (1994) examined children's speech samples using the CID Phonetic Inventory (Moog, 1989), which consists of a set of syllable imitation tasks in which the child is given points based on the number of contexts in which he/she produces particular sounds or suprasegmental skills correctly. Specifically, in terms of suprasegmentals, Tobey et al. judged imitations for correctness of duration, intensity, pitch, breath control, and overall voice quality in strings of one to three syllables. Each of these suprasegmental characteristics was given a score based on the number of contexts in which the characteristic attribute was produced correctly. The results reported by Tobey et al. showed that children improved in their suprasegmental production accuracy post-implantation (see also Kirk & Hill-Brown, 1985; Tobey, Angelette, Murchison, Nicosia, Sprague, Staller, Brimacombe, & Beiter, 1991; Tobey & Hasenstab, 1991; Tobey et al., 1994). Little is known, however, about the production of word-level suprasegmental characteristics, such as stress and syllable structure in these children.

Research with normal-hearing children has shown that an understanding of phonological development is incomplete without detailed investigation of the suprasegmental (i.e. prosodic) aspects of speech production. Between 1.5 and 4.0 years of age, young, normal-hearing children undergo a great deal of phonological development in their speech production skills. In the beginning, they tend to show poor segmental accuracy, often departing from the adult target model in substantial ways (Macken, 1980; Ingram, 1986; Menn & Matthei, 1992). However, it is often observed that they are able to reach adult-like word-level prosodic patterns prior to gaining accuracy on a segmental level (Menn, 1978; Gleitman & Wanner, 1982; Kirk & Hill-Brown, 1985; Stemberger, 1988; Echols, 1993; Peters & Menn, 1993).

One instance of children's accurate prosodic representations with incorrect segmental assignment is the case of segmental substitutions. For example, children's speech error patterns often consist of sound substitutions or sound exchange errors (e.g., 'pasghetti' for 'spaghetti') in which one or more segments are transposed or substituted for others. Another example of the dissociation between prosodic accuracy and segmental accuracy is children's use of filler syllables. Gleitman and Wanner (1982) discuss the early inclusion of stressless syllables in young children's speech, but note that they surface in an 'undifferentiated form' such as a schwa [ə] or other lax vowel. Examples include the production of 'report card' as [ə-poɹtkɑɹd] (Gleitman & Wanner, 1982) and 'you put' as [ʊ - pʊt] (Peters & Menn, 1993). In each of these cases, the prosodic structure of the target has been correctly reproduced, while the segmental content has been replaced with a simpler sequence of sounds.

These types of utterances, in which the prosodic envelope is retained while segmental composition is simplified, are often treated in terms of different frames, or tiers, of the phonological system. Garrett (1980; 1982), for instance, suggests that sound exchange errors and sound substitution errors are a result of the segmental composition of lexical items being mapped onto an independently represented prosodic frame. Thus, segmental material is reflected distinctly from the prosody in the output.

Other more linguistically oriented models suggest a separation of phonological tiers (Goldsmith, 1976; Menn, 1978; Stemberger, 1988; Echols, 1993). The proposal for a nonlinear relation between phonological levels (feet, syllables, segments, segmental features, etc.) was formalized into Autosegmental Theory by Goldsmith and developed to explain tonal shift and floating tones (in which vowels and their accompanying tones do not always correspond one-to-one). Autosegmental theory was extended to account for the dissociation between segmental and suprasegmental aspects of children's utterances by Menn (1978), Stemberger (1988), and Echols (1993). According to this approach, children's syllable structure is located on one tier and their segmental content on another, allowing omission of segments but retention of the syllable structure. For example, the notion of separate segmental and syllabic tiers led Echols to propose that young children's underlying representations may be fully specified at the level of the syllable tier, but only partially specified at the level of the segmental tier, in order to explain the phenomenon of substitutions such as [bænə] for 'banana', in which the onset of an omitted syllable is combined with the rime and coda of a produced syllable.

Investigation of the prosodic development of children with CIs may provide new insights into their phonological development and how their development compares to normal development processes. The current study was an attempt to further this long-term goal. We examined stress- and syllable-level imitation scores using utterances from a sample of experienced pediatric cochlear implant users. The utterances were obtained from a nonword repetition task. Typically, in a nonword repetition task the child is asked to listen to and immediately repeat back a phonologically permissible sound sequence that has no semantic content. The nonword repetition task is complex because it requires the child to successfully complete multiple auditory, cognitive, linguistic, and articulatory speech-motor processes, without relying on visual cues or previous experience with the stimulus tokens.

Nonword repetition tasks have been used successfully to study the speech production skills of children with various language-learning difficulties (e.g., Kamhi, Catts, Mauer, Apel, & Gentry, 1988; Gathercole & Baddeley, 1990; Edwards & Lahey, 1998; Botting & Conti-Ramsden, 2001; Roodenrys & Stokes, 2001). Although normally developing children have been studied as comparison groups for the children with language disorders (Dollaghan & Campbell, 1998; Weismer, Tomblin, Zhang, Buckwalter, Chynoweth, & Jones, 2000; Roodenrys & Stokes, 2001), the nonword repetition task has also been used with normally developing children in studies of vocabulary size, reading abilities, and phonological working memory (see Michas & Henry, 1994; Gathercole, 1995; Metsala, 1999; Adams & Gathercole, 2000, among others).

The version of the nonword repetition task we employed in this investigation was originally designed to study individual differences in phonological working memory in young normal-hearing children (Gathercole, Willis, Baddeley, & Emslie, 1994; Gathercole & Baddeley, 1996). We first adopted this procedure in our lab to study individual differences in phonological working memory of normal-hearing children (Carlson, Cleary, & Pisoni, 1998), and then extended its use to study the phonological working memory and the speech of deaf children with cochlear implants (Dillon & Cleary, 2000; Cleary, Dillon & Pisoni, 2002). In the present study, we were specifically interested in using the children's nonword repetition responses to investigate the prosodic characteristics of their speech production skills, that is, to explore their ability to imitate and reproduce the correct number of syllables and stress patterns in nonsense words.

Method

Participants

Twenty-four children who participated in either the 1999 or 2000 Central Institute for the Deaf (CID) 'Cochlear Implants and Education of the Deaf' project (see Geers, Nicholas, Tye-Murray, Uchanski, Brenner, Crosson, Davidson, Spehar, Torretta, Tobey, Sedey, & Strube, 1999) were included in this study. Table 1 shows a summary of the demographic information about the children. The group consisted of 15 males and nine females who ranged in age from 8.2 to 9.9 years (M = 8.8 years). Nineteen children were congenitally deaf; the other five children were three years old or younger at the onset of deafness. All of the children had used their cochlear implant for at least 3.8 years (M = 5.4 years), and all used a Nucleus 22 implant with the SPEAK coding strategy at the time of testing. Both oral and total communication modes were represented in the group.

Stimulus Materials

The stimuli used for this study were a subset of the 40 nonwords originally developed by Gathercole for the Children's Test of Nonword Repetition (CNRep; Gathercole et al., 1994; Gathercole & Baddeley, 1996). The specific items were selected by eliminating the CNRep nonwords that showed the least amount of variance in scores obtained previously in our lab from a young group of normal-hearing children (Carlson et al., 1998). We also eliminated nonwords that were essentially common real words attached in an unfamiliar manner to a standard affix. The remaining 20 nonword stimuli are shown in Table 2. These items were balanced in terms of syllable number: there were five words each, at syllable lengths of two, three, four, and five. Each of the target nonwords contained primary stress on either the first or second syllable.

Child ID #	Gender	Age	Age at Onset of Deafness	Duration of Deafness	Duration of CI Use
1	M	9.1	0	3.3	5.8
2	M	8.7	0	3.9	4.8
3	M	8.2	0	3.3	5.0
4	M	9.5	0	4.7	4.8
5	F	8.3	0	2.1	6.1
6	M	8.3	0	2.1	6.1
7	F	9.4	3.0	2.0	4.3
8	M	8.2	1.5	0.7	6.0
9	M	8.7	0	4.0	4.8
10	M	9.9	0	4.5	5.4
11	M	9.0	0	3.4	5.6
12	F	8.7	0	2.2	6.5
13	M	9.5	0	2.9	6.6
14	F	8.5	0	3.2	5.3
15	F	8.3	0	2.4	5.9
16	F	8.3	0	3.0	5.4
17	M	8.4	0	3.9	4.5
18	M	8.2	0.8	2.7	4.7
19	M	8.2	2.0	1.7	4.5
20	M	9.0	1.5	1.6	6.0
21	M	9.0	0	2.6	6.4
22	F	8.4	0	3.3	5.1
23	F	9.1	0	5.4	3.8
24	F	9.7	0	3.2	6.5
Means:		8.8	0.4	3.0	5.4

Table 1. Demographic information for the 24 children (ages and durations in years).

Procedure

Because the nonwords used by Gathercole et al. (1994) were originally recorded by a British talker, they were re-recorded in our lab by an adult female talker of American English (Carlson et al., 1998). The tokens were sampled at a rate of 22.05 kHz and stored as individual digital files. The stimuli were presented auditorily to the children via a desktop speaker (Cyber Acoustics MMS-1) at approximately 70 dB SPL. The children heard the list of nonwords in random order. They were forewarned that the stimuli would be unfamiliar 'funny' words and were told to imitate and reproduce the items to the best of their ability. Children's responses were recorded via a head-mounted microphone (Audio-Technica ATM75) onto digital audiotape (DAT) using a TEAC DA-P20 tape deck. The utterances on the DAT tapes were later digitized and segmented into individual sound files. Each imitation response was independently transcribed using broad phonemic transcription by the first and second authors. Intertranscriber agreement was 93%, and any disagreements about the transcriptions were resolved by a third transcriber.[2]

[2] We would like to thank Cynthia Clopper for her assistance as the third transcriber.

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 current study (adapted from Gathercole et al., 1994).

Scoring

Imitations were scored for accuracy in three ways. First, the responses were scored for overall phonological accuracy, where overall phonological accuracy meant the child's response was correct in terms of the reproduction of all segmental and suprasegmental features of the nonword target pattern. Second, the responses were scored for syllable accuracy, that is, whether the child correctly reproduced the same number of syllables as were present in the target nonword, regardless of segmental content. Third, the responses were scored for primary stress placement, that is, whether the child correctly reproduced the stress pattern that was present in the target. These latter two accuracy scores (the syllable score and the stress score), which assessed children's ability to correctly reproduce the prosodic characteristics of the target nonword, will be referred to throughout the remainder of the paper as the two prosodic accuracy scores.

Predictions

We predicted that children with cochlear implants would perform poorly on the overall phonological score. This was based on previous findings indicating difficulties in correctly reproducing segments and features of isolated syllables and words (Geers & Tobey, 1992; Sehgal, Kirk, Svirsky, Ertmer, & Osberger, 1998; Dillon & Cleary, 2000). However, as noted above, although young normal-hearing children may have some difficulty producing segmentally correct utterances, they are able to produce prosodic patterns that are faithful to the adult target (Menn, 1978; Gleitman & Wanner, 1982; Kirk & Hill-Brown, 1985; Stemberger, 1988; Echols, 1993; Peters & Menn, 1993). Because the deaf children in our study had three or more years of experience using their cochlear implant, we expected that they would possess a phonological system sufficient for them to produce nonword imitations that resembled the overall prosodic shape of the nonword targets, regardless of their accuracy on a segmental level. Specifically, we predicted that children's responses would be closer to the target with respect to the overall prosodic envelope than the detailed segmental pattern. Our expectation provided the motivation for the prosodic scoring system we developed.

In addition, we predicted that certain prosodic characteristics of the target patterns would affect children's performance on the two prosodic accuracy scores in specific ways. First, based on previous findings that overall repetition accuracy scores are higher for imitations of shorter target patterns (e.g., Gathercole, 1995), we predicted that the deaf children in our study would have higher prosodic accuracy scores for shorter nonwords than longer nonword patterns. Children should show higher syllable and stress scores in imitating a nonword like 'ballop' than a nonword such as 'detratapillic'.

Second, both prosodic accuracy scores (syllable and stress) should be higher for targets with primary stress on the initial syllable. The motivation for this prediction was based on earlier accounts of normal-hearing children's prosodic development (e.g., Demuth, 1995; Gerken, 1996), which suggest that young children produce words that are more prosodically faithful to the adult target if the target begins with a stressed syllable than an unstressed syllable. Thus, children should show higher prosodic scores for their imitation of the nonword 'fénnerizer' than for a nonword such as 'emplífervent'.

Finally, both prosodic accuracy scores should be higher for nonword targets with less syllable complexity (fewer consonant clusters, e.g., 'ballop') than patterns with higher complexity (e.g., 'prindle'). This prediction is based on the findings reported by Gathercole (1995), who showed that imitations of nonwords containing fewer consonant clusters tend to have greater overall accuracy. Therefore, we predicted that nonwords containing fewer consonant clusters would also display higher prosodic accuracy scores.

An additional set of analyses was made regarding the assessment of individual differences in the component processes of speech perception and production, including working memory, as reflected in the children's nonword repetition performance. Success in a nonword imitation task is contingent upon successful performance in encoding, storage in short-term memory, and production of a response. Based on recent findings demonstrating that nonword imitation ability correlates with performance on other speech- and language-related behavioral tests (Cleary et al., 2002), we predicted that higher prosodic accuracy scores (syllable and stress) would correlate with better performance on a range of speech perception and word recognition tests that were also collected from the 24 children during their participation in the CID summer programs. These behavioral tests included measures of spoken word recognition, language comprehension, speech intelligibility, speaking rate, and working memory span.

In addition, we also predicted that the syllable and stress scores obtained from the nonword repetition task would be related to a measure of overall perceived accuracy of the children's imitations that was obtained as part of a separate study with 240 native speakers of English who had no previous experience with deaf speech. In this study, ten listeners rated the imitations of each of the 24 children. For each rating, the listeners heard a 'model' target utterance followed by a child's imitation of that model, and gave a goodness rating of that child's imitation using a scale from '1' to '7', in which '1' represented a poor imitation of the model and '7' represented a perfectly accurate imitation. The listeners were only told to rate the child's imitation of the model utterance. They were not given any guidelines as to what to base their goodness ratings on, apart from being asked to ignore differences in pitch between the adult model's voice and the child's voice. We predicted that adults would judge children's imitations of the model target to be more accurate overall when the children's nonword responses had higher prosodic accuracy scores, regardless of the segmental accuracy of the children's utterances.

Results

Accuracy Scores

As predicted, the children performed quite poorly on the overall phonological accuracy score. Only 5% of their nonword imitations were produced correctly without any errors. However, further examination of the children's responses revealed that this overall score was based primarily on *segmental* errors. Importantly, children correctly reproduced target-like prosodic characteristics much more frequently than target-like segmental properties when the responses were scored for syllable number and stress placement. 48% of the nonword imitations were reproduced with both the correct syllable number and the correct primary stress placement. That is, almost half of the children's nonword repetitions were correct when these two prosodic dimensions were used to score the responses. Moreover, there were some imitations that were correct in terms of only one of the two prosodic characteristics. Of the total number of children's imitations, 64% of the responses were reproduced with the correct number of syllables as the target nonword, and 61% of the responses were reproduced with the correct placement of primary stress. Thus, analysis of the scores using these three accuracy measures indicate that these deaf children are able to reproduce the suprasegmental features of the target pattern much better than the segmental components of the nonword patterns.

Prosodic Analyses

Number of syllables in the target. Correlations were calculated between the number of syllables in the target pattern and the two prosodic accuracy scores described above (syllable score and stress score) for each of the 20 nonwords (averaged across children). The analysis revealed that the number of syllables in the target pattern was negatively correlated with syllable scores (r = -.25, p < .01). Likewise, the number of syllables in the target was also negatively correlated with stress scores (r = -.30, p < .01). Both correlations confirm our first prediction regarding the reproduction of the prosodic characteristics of the target pattern and indicate that children show more accurate performance on imitation of shorter target nonwords. Specifically, children preserved syllable number and primary stress location in their imitations more often for shorter nonword target patterns than for longer nonword patterns. The more syllables that were present in the target nonword, the less likely the children were to reproduce the correct number of syllables and the correct primary stress pattern of the target nonword.

In order to investigate the types of syllable errors that children made, their nonword imitations were broken down into the percent of nonword imitations that contained no errors, the percent of total imitations that contained syllable deletion errors, and the percent of total imitations that contained syllable addition errors. The results are shown in Figure 1. For targets with two, three, or four syllables, the majority of the imitations contained the correct number of syllables. For targets with five syllables, there were many fewer imitations that contained the correct number of syllables than imitations with syllable additions or deletions. Overall, when children failed to reproduce the correct number of syllables, they were more likely to delete syllables (70% of errors) than add syllables (30% of errors), except in the case of target nonwords with two syllables, for which only syllable additions were observed. Moreover, we also found that when a syllable error occurred, the response consisted of the addition or deletion of a single syllable (85% of errors). Only 14% of the syllable errors consisted of the addition or deletion of two syllables, and in one single case, the syllable error consisted of the addition of three syllables. Thus, shorter target nonwords resulted in fewer errors in syllable number. However, when children did produce errors, they tended to omit syllables more often than add syllables.

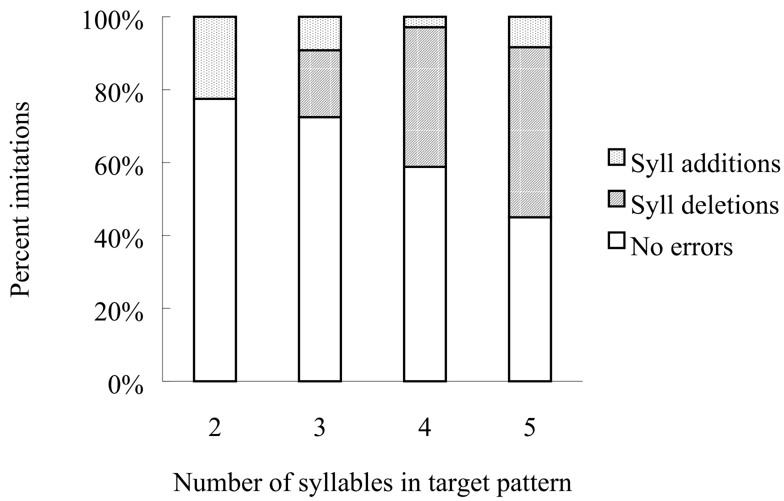


Figure 1. Percent of nonword imitation responses containing correct syllable count, syllable deletions, or syllable additions.

Stress placement in the target. To assess our second prediction that prosodic accuracy scores would be higher for stress-initial targets than non-stress-initial targets, we carried out two comparisons. The first was a comparison of the syllable scores for stress-initial targets and non-stress-initial targets. The second was a comparison of the stress scores for stress-initial targets and non-stress-initial targets. Figure 2 shows a summary of the results of these analyses. A significant difference in syllable scores between targets with primary stress on the initial syllable and targets with primary stress on the non-initial syllable was found, $t(478) = 2.7$, $p < .01$. A significant difference was also found for stress scores, $t(478) = 2.3$, $p < .05$. These two results show that children's imitations retained prosodic properties of targets with initial stress more often than targets with non-initial stress. If the target pattern had initial stress, the corresponding nonword imitations were more likely to contain the correct number of syllables and the correct location of primary stress.

Syllable complexity of the target. Our third prediction, that prosodic accuracy scores would be higher for targets containing fewer clusters, was not confirmed in our analyses. Specifically, neither the correlation between the number of consonant clusters in a target and syllable scores, nor the correlation between the number of clusters and stress scores reached significance. These findings indicate that children were able to reproduce prosodic characteristics equally well for nonwords with more consonant clusters than with fewer clusters, and that target syllable complexity did not affect the children's ability to reproduce prosodic structure in these patterns. This may be a reflection in the nonword patterns of a dissociation between syllable-level and segmental-level representations. The presence of consonant clusters may only affect children's performance on imitating the segmental components of those clusters, not children's performance on imitating and reproducing the more global overall prosodic shape of the syllables containing them.

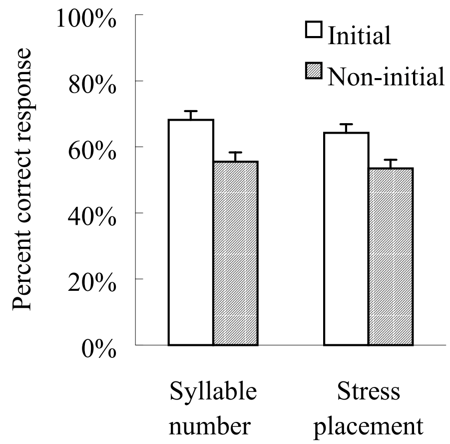


Figure 2. Mean percent correct as a function of target stress placement (initial syllable or non-initial syllable) and type of accuracy score (syllable or stress). Error bars show standard error.

Correlations with other Measures of Speech and Language Performance

We were also interested in the extent to which the children's prosodic scores on the nonword imitation task would reflect the contribution of the underlying component processes involved in speech perception, production and working memory. Although the nonword repetition task used in the present study may appear to be relatively simple at first glance, successful imitation of a nonword pattern involves the contribution of several different component processes: auditory and phonological encoding, short-term storage of the target item in working memory, and articulatory planning and speech production at the time of output. In order to be able to imitate and quickly reproduce a nonword pattern on the fly, a child needs to be able to perform well in each of these component processes.

Measure		Syllable score	Stress score
Word recognition	LNT Easy	+.38	+.57**
	LNT Hard	+.38	+.57**
	MLNT	+.42*	+.41*
	BKB	+.48*	+.48*
	WIPI	+.46*	+.36
Comprehension	TACL	+.68**	+.30*
Speech intelligibility	McGarr Sentences	+.50*	+.60*
Speaking rate	McGarr 7-syllable Sentences	-.55**	-.52**
Short-term memory	WISC (Forward Digit Span)	+.47*	+.57**
Executive function	WISC (Backward Digit Span)	+.14	+.18
Perceived Accuracy (Goodness)	Perceptual Ratings	+.67**	+.69**

** p < 0.05, ** p < 0.01*

Table 3. The correlation r-values between language measures, working memory measures, and perceptual ratings scores, and each of the prosodic accuracy scores.

The 24 children in this study also participated in a range of other tasks that were designed to measure their performance on these component processes as part of a larger study at CID (Geers et al., 1999). These scores provided an unusual opportunity to assess the contribution of several of these component processes to performance on the nonword repetition task. To accomplish this, correlations between the children's prosodic accuracy scores on the one hand, and several speech, language, and working memory measures on the other hand, were examined. Table 3 provides a summary of these correlations.

Correlations with word recognition measures. Three measures of spoken word recognition performance were available from CID: the Lexical Neighborhood Test (LNT), Bamford-Kowal-Bench Sentence List Test (BKB) and Word Intelligibility by Picture Identification test (WIPI). The LNT (Kirk, Pisoni, & Osberger, 1995) is an open-set test of spoken word identification consisting of 100 monosyllabic words divided into four lists of 25 words each. Two of the lists, the LNT Easy lists, contain words that are 'lexically easy' (i.e. phonetically similar to very few other words) and two of the lists, the LNT Hard lists, contain words that are 'lexically hard' (i.e. phonetically confusable with many other words). Each child was tested on one LNT Easy word list and one LNT Hard word list. Separate percent-words-correct scores were obtained for each list. Scores were also available from the Multisyllabic Lexical Neighborhood Test (MLNT), which is analogous to the LNT, but uses multisyllabic words of two or three syllables. The BKB is an open-set task involving spoken repetition of a target sentence (Bench, Kowal, & Bamford, 1979). The WIPI is a closed-set measure of spoken word identification involving a six-alternative pointing response (Ross & Lerman, 1979). The LNT, BKB, and WIPI were all administered using recorded auditory-only presentation.

Scores on LNT Easy words were positively correlated with the stress scores and syllable scores, but only the correlation between LNT Easy words and stress scores reached significance ($r = +.57$, $p < .01$). Likewise, scores on the LNT Hard words were positively correlated with the stress scores and syllable scores, but only the correlation between the LNT Hard words and stress scores reached significance ($r = +.57$, $p < .01$). Scores on the MLNT were significantly correlated with both the syllable scores and the stress scores ($r = +.42$, $p < .05$, $r = +.41$, $p < .05$, respectively). Likewise, BKB scores were significantly correlated with both of the prosodic accuracy scores (*r's* = +.48, $p < .05$). Lastly, the correlation between the children's WIPI scores and their syllable scores was significant ($r = +.46$, $p < .05$), although the correlation between the WIPI scores and their stress scores did not reach significance.

Taken together, these results show that children's ability to imitate the stress pattern and correctly reproduce the number of syllables in a novel nonsense word is related to their scores on several independent word recognition tasks. The findings suggest that the individual differences among the children in terms of their ability to imitate novel patterns is related to differences in their ability to recognize real words, both in open and closed set tasks as well as sentence contexts. The results suggest that the same underlying linguistic processes are being used in each of these component tasks to construct phonological representations.

Correlations with language comprehension. The battery of tests administered by researchers at CID also included the Test of Auditory Comprehension of Language Revised (TACL-R), a language comprehension measure that was designed to assess children's receptive vocabulary, morphology, and syntax (Carrow-Woolfolk, 1985). In the CID study, the TACL-R was administered using total communication to all children, and an age-equivalency score was obtained for each child. The TACL-R age-equivalent scores were positively correlated with both the children's syllable scores ($r = +.68$, $p < .01$) and their stress scores ($r = +.30$, $p < .05$). These results indicate that better performance on the nonword repetition task was associated with higher language comprehension scores in terms of receptive vocabulary, morphology, and syntax. That is, the children's ability to incorporate, reassemble and reproduce those specific phonological properties in their own speech production outputs is related to their ability to grasp regularities and patterns in their language environment.

Correlations with speech intelligibility. A measure of speech intelligibility was also obtained from each child as part of the larger study at CID using the McGarr Sentence Intelligibility Test (McGarr, 1983).[3] This test involves eliciting sentences three, five or seven syllables in length. Each child was provided with spoken and/or signed models of each test sentence as well as the printed text and was asked to repeat these sentences aloud as intelligibly as possible. The children's utterances were recorded and later played back to groups of adult naïve listeners who were asked to transcribe the utterances using standard orthography. This procedure provided an objective measure of speech intelligibility (McGarr, 1983).

Significant correlations were found between prosodic accuracy and speech intelligibility for both the syllable scores ($r = +.50$, $p < .05$) and the stress scores ($r = +.60$, $p < .05$), indicating that the children who produced more intelligible speech on the McGarr task also tended to correctly reproduce syllable number and primary stress location more often in their nonword imitations.

Correlations with speaking rate. Each child's productions of the McGarr sentences were also submitted to an acoustic analysis. One portion of the acoustic analysis consisted of measuring the duration of each seven-syllable sentence, and calculating an average sentence duration for each child.

Prosodic accuracy was negatively correlated with sentence duration. Specifically, mean sentence duration on the seven-syllable McGarr sentences was negatively correlated with syllable scores and stress scores ($r = -.55$, $p < .01$, $r = -.52$, $p < .01$ respectively). This result was not unexpected, as sentence duration is inversely related to speaking rate. The negative correlation between the prosodic accuracy scores on nonword repetition and speaking rate is consistent with previous findings showing that children who have slower speaking rates perform more poorly on speech production tasks (Cleary, Pisoni, Kirk, Geers, & Tobey, 2000; Pisoni and Cleary, in press). This finding suggests that limitations in phonological processing, coding, and rehearsal may be an important underlying factor leading to both slower speaking rate and poorer performance on a variety of speech and language measures.

Correlations with short-term memory. Measures of the children's forward and backward digit spans were also obtained for each child using the WISC Digit Span Supplementary Verbal sub-test of the Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991). The 'digits forward' task is considered a reliable measure of the encoding, rehearsal, and storage processes involved in short-term memory (Rosen & Engle, 1997; Pisoni & Geers, 2000; Engle, 2002). In contrast, the 'digits backward' task is thought to involve different and more complex processing abilities than the digits forward task, and is frequently used as a measure of 'controlled' cognitive processing, or 'executive function' because subjects have to consciously carry out operations or procedures that make demands on a limited capacity processing system (Rosen & Engle, 1997; Engle, 2002).

For the digits forward 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 allowed. The digits backward task is similar

---

[3] The McGarr speech intelligibility and duration measures were provided by Dr. Emily Tobey and her colleagues at Callier Advanced Hearing Research Center at the University of Texas, Dallas.

to the digits forward task. The only difference in procedure is that the child is asked to repeat the digits in reverse order from the order in which they were originally presented.

The children's WISC forward digit span scores were found to be positively correlated with both the syllable scores ($r = +.47$, $p < .05$) and the stress scores ($r = +.57$, $p < .01$). Longer forward digit spans were associated with higher scores on both of the prosodic measures of nonword repetition performance. The correlations between the children's WISC backward digit span scores and their prosodic scores were not significant (syllable scores, $r = +.14$; stress scores, $r = +.18$), indicating that backward digit span, unlike forward digit span, is unrelated to the ability to imitate prosodic characteristics of nonwords. This pattern is consistent with previous findings suggesting that forward and backward digit spans measure fundamentally different cognitive processes (Rosen & Engle, 1997; Engle, 2002).

Correlations with perceptual ratings. The nonword utterances produced by the children were also subjected to a perceptual analysis to obtain objective ratings of speech intelligibility and goodness judgments (for preliminary findings from a related study, see also Cleary et al., 2002). As described above, a perceptual measure of goodness was gathered from 240 adult listeners who heard model target utterances followed by the children's imitation responses and rated them for repetition accuracy on a scale from '1' to '7'. The ratings gathered from that study were averaged across listeners and test items to produce an average composite perceptual rating for each child. We then calculated correlations between these average perceptual ratings and our two prosodic accuracy scores (syllable score and stress score).

The perceptual ratings of the children's imitations were positively correlated with both of the prosodic scores (syllable score, $r = +.67$, $p < .01$ and stress score, $r = +.69$, $p < .01$) indicating that successful reproduction of prosody, that is, imitation of the correct number of syllables and the correct stress pattern of an utterance, is an important underlying phonological factor that influences adult listeners' perception of CI children's speech production accuracy in imitating nonword patterns (see also Cleary et al., 2002). These findings suggest that the children's ability to encode and reproduce the prosodic characteristics is linked to listeners' judgments of how well the speech production of deaf children with CIs resembles that of a normal hearing adult.

Discussion and Conclusions

In this investigation of the prosodic aspects of pediatric cochlear implant users' nonword imitation skills, we observed a number of related findings that provide some new insights into the underlying phonological skills that deaf children acquire after cochlear implantation. The children with CIs produced very few responses that matched the nonword target patterns exactly when their utterances were scored in terms of both segmental and prosodic attributes. Only 5% of their nonword imitations were correct by this strict scoring procedure. The children's performance was markedly lower than the typical performance of their normal-hearing peers who consistently perform near ceiling on the same nonword repetition task (Gathercole et al., 1994; see also Carlson et al., 1998). However, in light of previous studies showing poorer segmental accuracy in the speech production of children with cochlear implants, this finding was not unexpected. Although these deaf children had difficulty reproducing the segmental content of these nonword patterns, they displayed much greater skill in imitating suprasegmental properties of the nonwords. As expected, they were able to reproduce the correct number of syllables in the target pattern as well as the primary stress location on almost two-thirds of their nonword imitations. They also achieved significantly higher scores for targets with fewer syllables and with initial stress than for targets with more syllables and non-initial stress. Moreover, about half of the nonword responses contained *both* the correct syllable number and primary stress location, suggesting a knowledge of and sensitivity to select aspects of phonological structure in these nonword patterns.

The children's performance on the nonword repetition task suggests that they were able to acquire several important aspects of English prosodic structure and successfully access and use this prosodic knowledge in their utterances. The different speech production skills of these children at the suprasegmental and segmental levels lends further support for the proposal of a theoretical framework consisting of separate phonological tiers. It may be the case that during the process of encoding new words deaf children with CIs are more likely to correctly encode elements on the suprasegmental tier than the segmental tier, which requires encoding of much finer phonetic detail. Another possibility is that once the novel sound patterns are encoded, their output representations become more robustly specified at the level of the syllable than the segment. Further studies are necessary in order to make a particular claim as to where in the perception – production system these separate levels exist, but the present results clearly demonstrate a strong and reliable dissociation between segmental and suprasegmental levels in this task.

When children made errors on syllable number, their errors revealed a pattern that was similar to normal-hearing children. That is, normal-hearing children tend to delete syllables in their utterances more often than add them (Demuth, 1995). Taken together with the children's higher prosodic scores for targets with fewer syllables, this finding may reflect the additional load on short-term phonological memory capacity induced by a greater number of syllables in a temporal or segmental pattern, especially a nonword pattern that has no lexical representation. Further support for this conclusion was provided by the positive correlations observed between the children's prosodic scores and their forward digit spans. Longer digit spans were associated with higher stress and syllable scores, suggesting close links between processing capacity and encoding phonological structure (Gathercole, 1995).

The finding that prosodic performance was better for nonword targets with initial stress than non-initial stress is also consistent with previous findings regarding the English stress system in utterances of young normal-hearing children who are in the process of acquiring language (Demuth, 1996; Gerken, 1996). The majority of English content words have primary stress on the initial syllable (Cutler & Carter, 1987), thus the input that children perceive most often contains stress-initial forms. It is not surprising then that, when given a novel form that may tax the information processing system, children reproduce a stress pattern that is consistent with their prior experience and exposure to sound patterns in their ambient language.

The nonword imitations which erroneously contained a stress-initial syllable, whether it was formed by an omission of a stressless initial syllable, or by a stress-shift, are also similar to normal-hearing children's errors. These patterns are less faithful to the adult target but are more prosodically optimal, with regard to English stress (Demuth, 1995; 1996; Gerken, 1996). One account of this phenomenon, known as Optimality Theory (OT; Prince & Smolensky, 1993), assumes the existence of a set of ranked violable constraints in the child's phonological grammar, which yields his/her output forms. These constraints, while identical in type to those in an adult's grammar, are ranked in a different order for young children, and only over time do they adapt to reflect adult grammars (Demuth, 1995; Gerken, 1996; Pater & Paradis, 1996). At a young age, constraints that reflect common prosodic properties of the English language are ranked higher than constraints of segmental faithfulness to the adult target. This early ordering of constraints thus yields more prosodically optimal word forms in children, at the expense of segmental accuracy. Such an account may be useful for the productions in the current study, although an in-depth discussion of constraints and OT is outside the purview of the present report.

The finding that the complexity of consonant clusters did not affect syllable number or stress accuracy scores is inconsistent with the earlier findings of Gathercole (1995), who reported an effect of consonant clusters on segmental accuracy. However, the present results can be explained by considering differences in the perception of segmental and suprasegmental properties of these nonword patterns: the presence of consonant clusters may only affect how well children can perceive and imitate the constituents of those clusters, not how well children can imitate the prosodic shape of the syllables containing them. It is well-known that deaf children with CIs have difficulty perceiving fine phonetic distinctions such as place of articulation and voicing in stop consonants (Miyamoto, Kirk, Todd, Robbins, & Osberger, 1995; Chin & Finnegan, 1998).

The correlations between the prosodic scores, and word recognition as well as the language comprehension scores suggest that the underlying linguistic processes that enable a deaf child with a cochlear implant to accurately imitate the prosodic structure of novel patterns are related to his/her real-word recognition and comprehension skills. The presence of correlations between the children's prosodic scores and both the McGarr intelligibility measure and the goodness ratings by naïve adults indicate that preservation of prosody may be an important factor affecting adult listeners' perception of deaf children's speech intelligibility. The correlations between the prosodic accuracy scores and the speaking rate measure are consistent with previous findings showing that children with faster speaking rates perform better on a range of speech production tasks. Specifically, Pisoni and Cleary (in press) found a positive correlation between speaking rate (as measured by McGarr durations) and WIPI scores, LNT scores, BKB scores, and digit spans for 88 children who participated in the 1998 and 1999 summer programs at CID. Our results showing that children who had higher prosodic accuracy scores in the nonword repetition task also had higher word recognition scores, higher language comprehension scores, slower speaking rates and longer digit spans extend and refine the earlier findings of Pisoni and Cleary (in press). Together, these results suggest that a common underlying source of variance related to phonological processing skills is operative and that these fundamental skills influence children's performance not only on prosodic accuracy of nonword imitations but also on all of these other behavioral tasks.

Taken together, the results of this study demonstrate that the nonword repetition task and the component information processing and linguistic processing that it taps into can provide new insights into the speech production skills and underlying linguistic abilities of deaf children following cochlear implantation. The present findings indicate that experienced pediatric CI users are able to encode the prosodic structure of nonwords that conform to English phonological rules. They are able to reproduce syllable and stress information with relatively high levels of accuracy, despite their difficulty in perceiving and reproducing the fine segmental properties of these novel patterns. The present findings also demonstrate a close correspondence between the children's speech perception, working memory and speech production skills. With further analytic studies of this type using novel information processing tasks, we hope to better understand the relationships between perceptual, cognitive, and linguistic skills used in the processing of spoken language, and describe how these fundamental information processing skills develop and change over time in deaf children following cochlear implantation.

References

Adams, A. & Gathercole, S.E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language and Communication Disorders, 35,* 95-116.

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.

Botting, N. & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *International Journal of Language and Communication Disorders*, *36*, 421-432.

Brown, C. & McDowall, D.W. (1999). Speech production results in children implanted with the CLARION implant. *Annals of Otology, Rhinology, and Laryngology (Suppl.)*, *17*, 7110-7112.

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.

Chin, S.B. & Finnegan, K.R. (1998). Minimal pairs in the perception and production of speech by pediatric cochlear implant users: A first report. In *Research on Spoken Language Processing Progress Report No. 22,* (pp. 291-303). Bloomington, IN: Speech Research Laboratory, Indiana University.

Cleary, M., Dillon, C. & Pisoni, D.B. (2002). Imitation of nonwords by deaf children after cochlear implantation: Preliminary findings. *Annals of Otology, Rhinology and Laryngology (Suppl., 111,* 91-96.

Cleary, M., Pisoni, D.B., Kirk, K.I., Geers, A., & Tobey, E. (2000). Working memory and language development in children with cochlear implants. Poster presented at CI 2000: The 6th International Cochlear Implant Conference. February 2000, Miami, Florida.

Cutler, A. & Carter, D.M. (1987). The predominance of strong initial syllables in the english vocabulary. *Computer Speech and Language*, *2*, 133-142.

Demuth, K. (1995). Markedness and the development of prosodic structure. In: J. Beckman (Ed.) *Proceedings of NELS 25,* (pp. 13-25). Amherst, MA: GLSA, University of Massachusetts.

Demuth, K. (1996). Prosodic structure of early words. In: J.L. Morgan & K. Demuth (Eds.) *Signal to Syntax: Boot-strapping from Speech to Grammar in Early Acquisition*, (pp. 171-184). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

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

Echols, C. (1993). A perceptually-based model of children's earliest productions. *Cognition*, *46*, 245-296.

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.

Garrett, M.F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language Production Volume 1: Speech and Talk* (pp. 177-220). London: Academic Press.

Garrett, M.F. (1982). A perspective on research in language production. In J. Mehler, E.C.T. Walker & M. Garrett (Eds.), *Perspectives on Mental Representation: Experimental and Theoretical Studies of Cognitive Processes and Capacities* (pp. 185-199). Hillsdale, NJ: Lawrence Erlbaum Associates.

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. & 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. & 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. & 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.

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. *Central Institute for the Deaf Research Periodic Progress Report No. 35*, (pp. 5-20). St. Louis, MO: Central Institute for the Deaf.

Gerken, L. (1996). Prosodic structure in young children's language production. *Language*, *72*, 683-712.

Gleitman L. & Wanner, E. (1982). The state of the state of the art. In: L. Gleitman & E. Wanner (Eds.), *Language Acquisition: The State of the Art* (pp. 3-48). Cambridge: Cambridge University Press.

Goldsmith, J. (1976). *Autosegmental Phonology*. Cambridge, MA: Doctoral dissertation, MIT. Distributed by Indiana University Linguistics Club. New York: Garland Press, 1979.

Ingram, D. (1986). Phonological development: Production. In: P. Fletcher & M. Garman (Eds.), *Language Acquisition, Second Edition* (pp. 223-239). Cambridge: Cambridge University Press.

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

Kirk, K.I. & Hill-Brown, C. (1985). Speech and language results in children with a cochlear implant. *Ear and Hearing*, *6 (Suppl.),* 36-47.

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

Macken, M.A. (1980). The child's lexical representation: The 'puzzle-puddle-pickle' evidence. *Journal of Linguistics*, *16*, 1-17.

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

Menn, L. (1978). Phonological units in beginning speech. In A. Bell & J. Hooper (Eds.), *Syllables and Segments* (pp. 157-171). Amsterdam: North Holland Publishing Company.

Menn, L. & Matthei, E. (1992). The 'two-lexicon' account of child phonology: Looking back, looking ahead. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 211-247). Parkton, MD: York Press.

Metsala, J. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*, 3-19.

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

Miyamoto, R.T., Kirk, K.I., Todd, S.L., Robbins, A.M., & Osberger, M.J. (1995). Speech perception skills of children with multichannel cochlear implants or hearing aids. *The Annals of Otology, Rhinology, and Laryngology*, *104*, 334-337.

Moog, J.S. (1989). *The CID Phonetic Inventory*. St. Louis: Central Institute for the Deaf.

Pater, J. & Paradis, J. (1996). Truncation without templates in child phonology. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (Eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development* (pp. 540-551). Somerville, MA: Cascadilla Press.

Peters, A. & Menn, L. (1993). False starts and filler syllables: Ways to learn grammatical morphemes. *Language*, *69*, 742-777.

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.), *Handbook of Auditory Research: Auditory Prostheses, Springer Handbook of Auditory Research*. New York: Springer-Verlag.

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 (Suppl.), 18*, 592-593.

Prince, A. & Smolensky, P. (1993). Optimality Theory: Constraint interaction in generative grammar. *Technical Report #2*, Rutgers University Center for Cognitive Science. Cambridge, MA: MIT Press.

Roodenrys, S. & Stokes, J. (2001). Serial recall and nonword repetition in reading disabled children. *Reading and Writing*, *14*, 379-394.

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.

Sehgal, S.T., Kirk, K.I., Svirsky, M., Ertmer, D.J., & Osberger, M.J. (1998). Imitative consonant feature production by children with multichannel sensory aids. *Ear and Hearing*, *19*, 72-84.

Serry, T.A. & Blamey, P.J. (1999). A 4-year investigation into phonetic inventory development in young cochlear implant users. *Journal of Speech, Language, and Hearing Research*, *42*, 141-154.

Serry, T., Blamey, P., & Grogan, M. (1997). Phoneme acquisition in the first four years of implant use. *The American Journal of Otology*, *18 (Suppl.),* S122-S124.

Stemberger, J. (1988). Between-word processes in child phonology. *Journal of Child Language*, *15*, 39-61.

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 and Hearing*, *12 (Suppl.),* 48-54.

Tye-Murray, N., Spencer, L., Bedia, E.G., & Woodworth, G. (1996). Differences in children's sound production 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.

Weismer, S.E., 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.

Perceptual Ratings of Nonword Repetitions by Deaf Children after Cochlear Implantation: Correlations with Measures of Speech, Language and Working Memory[1]

Caitlin M. Dillon, Rose A. Burkholder, Miranda Cleary and David B. Pisoni

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

[1] This work was supported by NIH Research Grant DC00111 and Training Grant DC00012 to Indiana University. We would like to extend our thanks to Dr. Ann Geers and the research staff at Central Institute of the Deaf for testing the cochlear implant users and making the data available to us for this study. We are also grateful to Dr. Emily Tobey and her research team at Callier Advanced Hearing Research Center at the University of Texas, Dallas, for providing the McGarr intelligibility and duration data.

Perceptual Ratings of Nonword Repetitions by Deaf Children after Cochlear Implantation: Correlations with Measures of Speech, Language and Working Memory

Abstract. Seventy-six profoundly deaf children with cochlear implants completed a nonword repetition task. The children were presented with 20 nonword auditory patterns over a loudspeaker and were asked to repeat them aloud to the experimenter. All of the children were deaf before the age of 36 months. At the time of testing the children were between 8 and 9 years old. All but three of the children used a Nucleus 22 implant with the SPEAK coding strategy at the time of testing. Two of the children used a Clarion device and one child used a Nucleus 24 implant at the time of testing. The duration of implant use for all children was between 4 and 7.5 years. All of the children in this study produced a repetition response to at least 15 of the 20 target nonwords. The children's responses were recorded on digital audiotape and then played back to 240 normal-hearing adult listeners for accuracy judgments. Normal-hearing listeners rated the accuracy of the children's imitation responses against the original target models using a 7-point scale. The perceptual ratings of the children's nonwords were strongly correlated with scores on separate independent measures of spoken word recognition, immediate memory span, and verbal rehearsal processes used in phonological working memory, as well as speech production skills. Children who had become deaf at slightly older ages and children who had been deaf for shorter periods of time prior to implantation received higher perceptual ratings. Children whose early linguistic experience and educational environments emphasized oral communication methods also received higher ratings than children enrolled in total communication programs. The findings from this study suggest that individual differences in performance on the nonword repetition task are strongly related to variability observed in the component processes involved in nonword repetition including speech perception, encoding and verbal rehearsal in phonological working memory, and speech production. In addition, a shorter period of deafness prior to implantation and an educational environment emphasizing oral communication may be beneficial to deaf children's ability to develop the robust phonological processing skills that are necessary to accurately repeat novel, nonword stimuli. These skills appear to be related to more complex cognitive processes involved in word learning, vocabulary development, and speech production.

Introduction

The investigation of individual differences in the speech and language skills of deaf children after cochlear implantation has been a major focus of the research program in our laboratory for several years (Pisoni, Cleary, Geers, & Tobey, 2000). We are interested in how the variability in a range of speech and language outcome measures can be accounted for by differences in basic underlying cognitive processing skills such as phonological encoding, short-term storage, verbal rehearsal processes in phonological working memory, and speech production. These basic cognitive processes are assumed to be involved, in one way or another, in all behavioral tasks that require the immediate repetition of an auditory pattern such as a familiar word or a novel phonological nonword pattern.

The nonword repetition task in which a child is asked to immediately repeat back a novel stimulus pattern has been used extensively in recent years by researchers to study the development of phonological working memory in normal-hearing children (e.g., Gathercole, Willis, Baddeley, & Emslie, 1994). More recently, we have also used this methodology to study deaf children with cochlear implants (Cleary, Dillon, & Pisoni, 2002; Dillon & Cleary, 2000). In the present paper, we report data obtained from a nonword repetition task completed by 76 deaf children who use cochlear implants. The results provide several new insights into the relationship between individual variability in the component processes of phonological encoding, storage and speech production, and several traditional outcome measures of the benefit received from a cochlear implant.

Preliminary findings obtained from a small group of children were reported recently in Cleary et al. (2002). In that study, 14 children who had used cochlear implants (Nucleus 22, SPEAK coding strategy) for at least 3.5 years (*M* = 5.5 years) completed a nonword repetition task. Each of their responses was then rated for accuracy by 10 normal-hearing adult listeners on a scale from 1 to 7. We found that the mean ratings received by each child were strongly correlated with other independent measures of word recognition, language comprehension, working memory, speech intelligibility, and speaking rate. These results demonstrate that deaf children with sufficient experience with CIs are able to complete a complex task such as nonword repetition. In addition, the results indicate that the children's performance on this nonword repetition task may predict, at least in part, their performance on other traditional speech and language outcome measures.

The present study addressed the following three questions: First, to what extent is individual variability on the nonword repetition task observed in deaf children after cochlear implantation? Second, is individual variability among children in nonword repetition performance related to individual differences in each of the component processes of speech perception, working memory, verbal rehearsal, and speech production? Third, are the children's demographic characteristics related to their nonword repetition performance? Thus, the primary goal of this study was to replicate our earlier findings by collecting additional perceptual ratings of the nonword repetitions produced by a larger number of deaf children using cochlear implants. The CI users in the original study by Cleary et al. (2002) had all produced a response to each nonword stimulus. The children in the present study included the 14 children in Cleary et al. (2002), and 62 additional children who were somewhat more variable in their response output in the nonword repetition task. The children in the present study produced responses on 75 to 100% of the nonword stimuli. Other than these differences, the procedures were the same.

Methods

Participants

Seventy-six deaf children who use cochlear implants were studied. All of the children were tested at Central Institute for the Deaf (CID) in St. Louis, Missouri in 1999 or 2000 as part of a larger study called "Cochlear Implants and Education of the Deaf Child" conducted by Ann Geers and her colleagues (see Geers et al., 1999). Thirty-eight of the children were participants in 1999 and 39 in 2000. Prior to their inclusion in the larger study at CID, the deaf children were evaluated through intelligence testing to ensure that they fell within a reasonable range of intelligence expected for their chronological age. Only children that met this criterion were tested and included in the present study. Thirty-six of the children were female, and 40 were male.

Table 1 shows a summary of the demographic characteristics of the participants in this study. The mean age at onset of deafness for all children was just over two months (*M* = 2.3, *SD* = 6.4). Most of the children had a congenital profound hearing loss. Only 12 of the children lost their hearing after birth, between the ages of 1 and 36 months (*M* = 14.3*, SD* = 9.7). The duration of deafness before implantation was between 7 and 65 months (*M* = 37.2*, SD* = 13.1). The children's age at implantation was between 1.9 and 5.4 years (*M* = 3.3*, SD* = 1.0). All of the children had used their implants for 3.8 to 7.5 years (*M* =

5.6, *SD* = 0.8), prior to testing. At the time of testing, the children's ages ranged from 7.9 to 9.9 years (*M* = 8.9, *SD* = 0.6). Most children were using a Nucleus-22 device with the spectral peak (SPEAK) coding strategy at the time of testing. Several other children had either a Nucleus-24 implant or a Clarion cochlear implant. The number of active electrodes in the devices ranged from 8 to 22 (*M* = 18.4, *SD* = 2.3).

Demographic Variables	***Range***	***Mean***	***Standard Deviation***
Age at Onset of Deafness (in months)	0 - 36	2.3	6.4
Duration of Deafness (in months)	7 - 65	37.2	13.1
Age at Implantation (in years)	1.9 - 5.4	3.3	1.0
Duration of Implant Use (in years)	3.8 - 7.5	5.6	0.8
Chronological Age (in years)	7.9 - 9.9	8.9	0.6
Number of Active Electrodes	8 - 22	18.4	2.3
Communication Mode Score	6 - 30	19.8	7.7

Table 1. Summary of the demographic makeup of the 76 children.

The children were classified into two different groups based on whether they used primarily oral communication (OC) or total communication (TC) strategies. TC is a method utilizing manual signs and lip reading strategies, in addition to speech. The classification was based on scores assigned at five different points in time to evaluate the nature of the children's early experience and communication training programs. The five evaluations were made at: (1) pre-implantation, (2) one year post-implantation, (3) two years post-implantation, (4) three years post-implantation, and (5) at the time of testing (which was the time of their participation in the CID summer program). Scores for each evaluation ranged from "1," signifying a program stressing the use of signs (usually in the form of Signed Exact English (SEE), not American Sign Language (ASL)) and lipreading strategies to "6," representing an oral-only regime. The scores assigned at each evaluation were then summed to determine the communication mode of the CI children at the time of testing. The summed communication mode scores obtained for these 76 children ranged from 6 to 30 (*M* = 19.8, *SD* = 7.7). Children with communication scores of 15 and below were considered to be TC users. Children with communication scores above 15 were considered to be OC users. This classification was made based on the original scoring method in which the lower scores (1-3) most accurately represent TC methods, while higher scores (4-6) most accurately represent oral methods. Twenty-nine of the original 76 children were classified as TC, while the remaining 47 were classified as OC users.

Stimulus Materials

The children were tested with a shortened version of the Children's Test of Nonword Repetition (CNRep; Gathercole et al., 1994; Gathercole & Baddeley, 1996). This nonword repetition test was originally designed to evaluate the phonological working memory skills of normal-hearing and typically developing children. The 40 original nonwords in the CNRep are all phonotactically permissible patterns that could be real words in English. The test items used in the present study included a subset of 20 nonwords, selected from the 40 stimuli originally developed by Gathercole et al. (1994). These 20 stimuli were chosen because they elicited more variability in performance scores when tested on a sample of young, normal-hearing children (Carlson, Cleary, & Pisoni, 1998). The 20 stimuli are listed in Table 2, along with their phonetic transcriptions. The final set of 20 nonwords included five words each at syllable lengths of 2, 3, 4, and 5. The full set of stimuli for the CNRep was previously recorded in our laboratory by a female speaker of American English.

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

Table 2. The 20 nonwords used in the current study (adapted from Gathercole et al., 1994).

Procedure

Before any of the nonwords were presented, the children were told that they would hear some “funny, made-up sounding” words. They were also told that they would have to repeat each stimulus clearly out loud after they listened to it. The nonwords were then presented in random order to the children, one at a time, at approximately 70 db SPL through a tabletop loudspeaker (Cyber Acoustics MMS-1). The speaker was directly facing the children at a distance of approximately one and a half feet. If a child made the request, the sound level was increased to a comfortable level.

All of the children included in this study produced a verbal response to at least 15 of the 20 nonword stimuli. Figure 1 shows the frequency distribution of the number of nonword repetition responses completed by the children. Of the 76 children, 24 produced a response to each of the 20 target stimuli, 19 children produced 19 responses, 12 children produced 18 responses, 10 children produced 17 responses, 3 children produced 16 responses, and 8 children produced 15 responses.

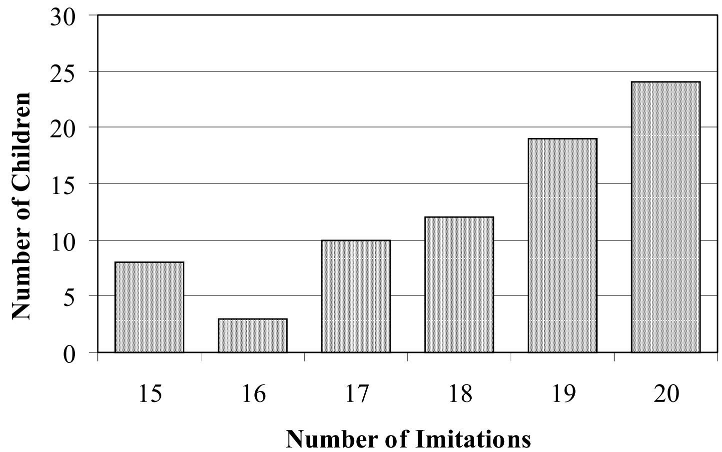


Figure 1. Frequency distribution of the number of nonword imitation responses produced by the 76 children (M = 18.36, SD = 1.63).

The children's nonword repetition responses were recorded onto digital audiotape (DAT; Sony Walkman TCD-D8) via a uni-directional headset cardioid condenser microphone (Audio-Technica ATM75). This equipment did not physically or mechanically interfere with the function or placement of the cochlear implant. The audio recordings were then digitized and transferred to computer files using the CoolEdit 96 (Syntrillium Software Corporation, 1996) digital waveform editing program. Each utterance was stored as an individual sound file.

All utterances were then edited and compiled together to make 24 different stimulus lists. The test lists were designed so that each list included the entire recorded data set of responses obtained from six different children. Each list included all of the utterances of one child who had produced a response to all 20 of the nonwords. Additionally, the data from five more children who produced a response to fewer than 20 stimuli were included to complete each list. The additional children assigned to each list had produced either 19, 18, 17, 16, or 15 utterances. By this method, no two children who had completed the same number of repetitions were included in a single test list. Thus, a complete list included 105 nonword repetitions produced by six children.

Given the distribution of the number of responses per child, it was necessary to include some children who did not complete all 20 utterances in more than one list so that every list would include 105 utterances. The lists were balanced in this way to insure that the adult listeners heard responses from several children whose set of imitation responses ranged from 15 to 20 utterances. In our earlier study, only children who had produced a response to all 20 of the stimuli were included (Cleary et al., 2002).

The test lists were then played back to normal-hearing adults who were asked to judge the accuracy of the nonword repetition responses produced by the children. Two-hundred-forty normal-hearing adult listeners, obtained from the Indiana University Psychology 101 Subject Pool, were asked to provide accuracy ratings for the CI children's imitations of the nonwords. All participants were monolingual native speakers of American English. They reported no history of a speech or hearing disorder at the time of testing. In addition, all subjects reported having no prior experience listening to the speech of the deaf or hearing impaired. After completing the perceptual rating task, the subjects received credit for the undergraduate psychology course in which they were currently enrolled.

Ten subjects heard each of the 24 test lists, so that each child was rated by at least 10 listeners. On each test trial, the listener first heard the target model stimulus pattern that had been presented to the deaf children for repetition. After hearing the model, the listeners then heard the imitation of that nonword stimulus produced by one of the children. The model stimulus and the child's response were separated by a one second interval of silence.

The order of presentation of the test trials was randomized within a list. Each listener completed two testing blocks using the same set of stimuli. Each block was randomized. Using this procedure, each nonword repetition was rated twice within the experiment by the same listener. Perceptual ratings were given according to a 7-point Likert scale between 1 and 7. The bottom of the scale, "1," was meant to represent a nonword repetition that "totally fails to resemble the target utterance." A score of "7" corresponded to nonword repetitions that were considered by the listeners to be "perfectly accurate renderings of the target utterance, ignoring differences in pitch." Listeners pressed buttons on a seven-button response box to record their judgments. The response box was interfaced to a PC that presented the stimuli to the listeners over high quality headphones (Beyer Dynamic, DT100) and recorded their responses. No feedback was provided during the experiment.

In addition to providing nonword repetition responses for this study, the children also completed a range of other tests designed to assess their speech perception, language, and literacy skills. To evaluate their spoken language comprehension skills, the Test of Auditory Comprehension of Language - Revised (TACL-R; Carrow-Woolfolk, 1985) was given. To evaluate their open-set speech perception and word identification skills, the children were tested using the Lexical Neighborhood Test (LNT; Kirk, Pisoni, & Osberger, 1995). Word identification skills were also measured through open-set sentence repetition using the Bamford-Kowal-Bench (BKB) test (Bench, Kowal, & Bamford, 1979). The Word Intelligibility by Picture Identification test (WIPI; Ross & Lerman, 1979) was used to measure closed-set word recognition skills. Speech feature discrimination was measured using the VIDSPAC (Boothroyd, 1997). The VIDSPAC is a video game test of speech feature contrast perception that was specifically designed for use with hearing-impaired children.

In addition to these traditional outcome measures, forward and backward digit spans were obtained using the Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991) to obtain estimates of the capacity of immediate memory. Forward digit span was used as a measure of the children's phonological working memory (Pisoni & Geers, 2000). Backward digit span was used to measure more complex, controlled cognitive processing abilities, or 'executive function' (see Engle, 2002; Lezak, 1983; Li & Lewandowsky, 1995; Rosen & Engle, 1997).

Two speech production measures were also available from another component of the larger CID project. Speech intelligibility scores were obtained using the McGarr Sentence Intelligibility Test (McGarr, 1981). This test contains a set of 3-, 5-, and 7-syllable sentences that are used to elicit spoken utterances from the children. These speech samples were later played back to normal-hearing adult listeners who were asked to transcribe the sentences. These transcription scores provide an objective measure of speech intelligibility. Duration measurements of these test sentences were also used to provide estimates of speaking rate. The duration measurements and the speech intelligibility scores for the McGarr sentences were made by researchers at the Callier Advanced Hearing Research Center at the University of Texas, Dallas, under the direction of Dr. Emily Tobey.

Results

Perceptual Ratings. As mentioned earlier, all of the listeners rated each stimulus item twice during the experiment, once during the first block of test trials and a second time during an additional block. A comparison of the accuracy ratings obtained in each block provided a way to assess the reliability of the perceptual ratings within listener. In order to compare performance in the two blocks, all of the nonword ratings provided by an individual listener in each block were averaged to compute the mean rating provided by each listener per block. The mean ratings in the first block (M = 2.96, SD = .48) were highly correlated (r = +.87, p <.001) with the mean ratings in the second block (M = 2.98, SD = .55). The difference between blocks was not significant.

We also calculated the mean rating per child for each block by averaging the ratings of all imitations produced by an individual child. The mean rating per child in the first block (M = 3.05, SD = 1.12) was also strongly correlated (r = +.99, p < .001) with the mean rating per child in the second block (M = 3.07, SD = 1.07). The mean rating per child for the first block was not significantly different from the mean rating for the second block. Because the perceptual ratings in block one were not different from the ratings in block two, when averaged across either listener or child, the following results are based only on the accuracy ratings provided by the listeners during the first block of trials.

As in our earlier report of nonword repetition in deaf children using cochlear implants (Cleary et al., 2002), we found a wide range of variability in the mean accuracy ratings given to the children by the group of normal-hearing listeners. The distribution of the mean ratings scores assigned to the 76 children is shown in Figure 2. Very few children received ratings at floor on this task, nor did any children receive ratings at the ceiling level, indicating that this perceptual rating task provided a sensitive measure of nonword repetition performance for this particular group of children.

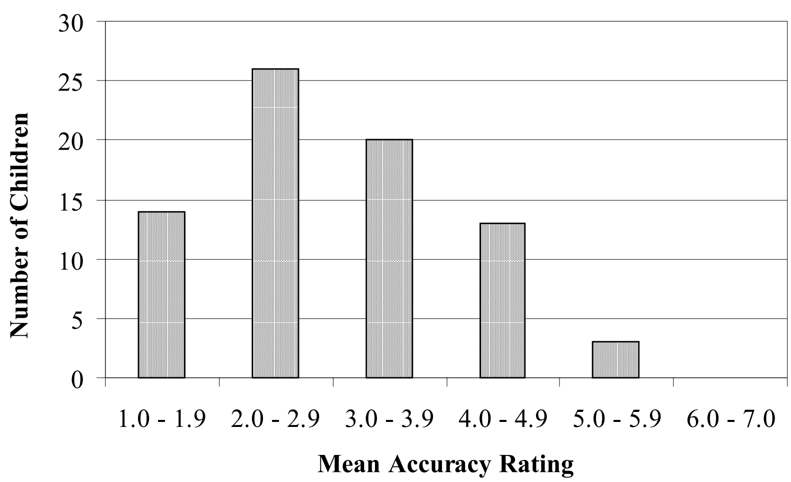


Figure 2. Frequency distribution of mean perceptual rating assigned to each child (N = 76) on a scale of “1” to “7” (M = 3.05, SD = 1.12)

Figure 3 shows a summary of the mean perceptual ratings of the OC and TC children’s nonword imitations. When the mean perceptual ratings of the children’s nonword imitations were summarized and divided into two separate groups, according to communication mode, a t-test revealed that the nonword responses from the OC children (M = 2.46, SD = .96) were rated significantly higher ($t(74)$ = -4.49, p <.001) than the nonword responses from the TC children (M = 1.40, SD = 1.05). This finding demonstrates an effect of early auditory and linguistic experience on nonword repetition performance. The children’s ability to reproduce a novel nonword stimulus pattern is better when, over time, the use of oral language has been encouraged over the use of manual sign and lip reading in their educational environment.

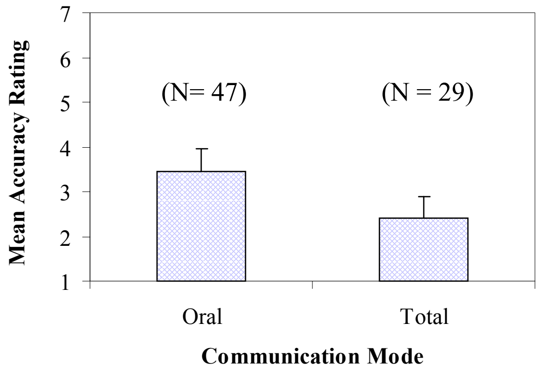


Figure 3. Mean perceptual rating of nonword imitations by OC and TC children. Error bars represent the standard error.

The mean perceptual rating for each of the 20 nonword patterns was also calculated based on the individual ratings of each response to that nonword. Figure 4 shows the mean perceptual ratings of the all the nonwords. The imitations of 2- and 3-syllable nonwords received higher ratings than the imitations of 4- and 5-syllable nonwords. An Analysis of Variance (ANOVA) revealed a significant main effect of syllable length on perceptual rating score ($F(19) = 4.38$, $p < .05$). Post-hoc Tukey tests revealed that imitation responses to 2- and 3-syllable target nonwords were each rated significantly higher than imitation responses to 5-syllable target nonwords (*p*'s = .05). This pattern may reflect the linguistically poorer imitations of the longer nonwords caused by their greater demands on phonological working memory (see also Carter, Dillon, & Pisoni, in press). In addition, the pattern of results shown in Figure 4 is consistent with findings that normal-hearing children's ability to accurately imitate shorter nonwords exceeds their ability to accurately imitate longer nonwords (Gathercole, 1995).

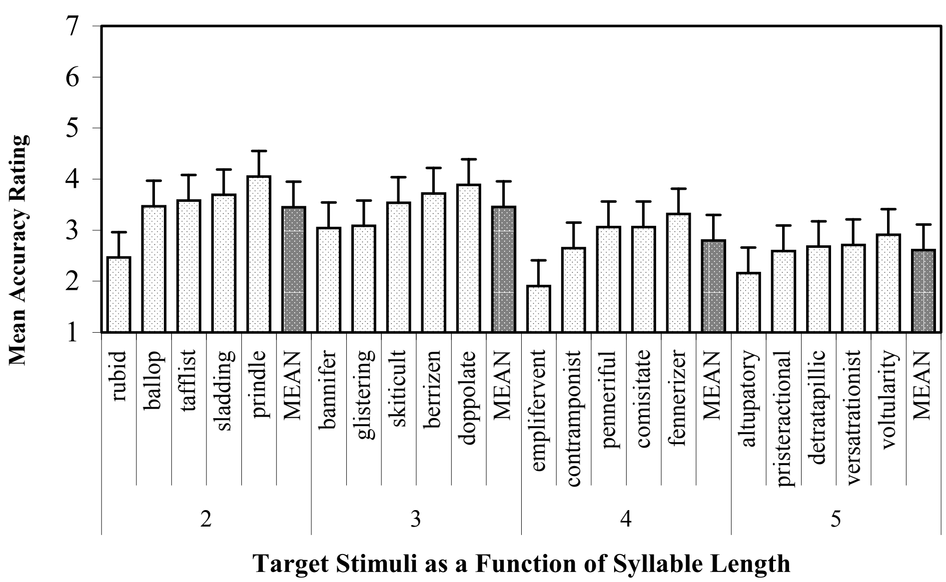


Figure 4. Mean perceptual ratings of each of the 20 nonwords, and the mean perceptual rating of each set of nonwords at a particular syllable length (2, 3, 4 or 5). Error bars represent the standard error.

Correlational Analyses. In order to investigate the extent to which the children's nonword repetition responses were related to their demographic characteristics and their scores on other measures of phonological encoding, immediate memory, verbal rehearsal, and speech production, we carried out a set of correlational analyses. A summary of the demographic correlations is shown in Table 3. The duration of time that the children had used their implants and their age at the time of testing were not significantly correlated with the nonword perceptual ratings. This is not surprising because these two factors were controlled in this study and the group was relatively homogeneous in terms of these two variables. Age at implantation was also not correlated with the perceptual ratings. The lack of correlation could indicate that there is no significant benefit of earlier implantation, or simply that children implanted at a later age had 'caught up' with the earlier-implanted children after four to five years of implant use. The absence of a correlation between age at implantation and nonword repetition performance could also be due to the relative homogeneity of this group of children who were part of the original experimental design.

Demographic Variables	Correlation *r* values
Age at onset of deafness	+.32**
Duration of deafness	- .26*
Age at implantation	- .12
Duration of cochlear implant use	+.06
Age at time of testing	- .13
Degree of exposure to oral-only communication	+.51***
Number of active electrodes in implant	- .31
Gender	- .14

Table 3. Correlations between nonword repetition perceptual ratings and demographic variables of the deaf children using cochlear implants. *$p < .05$, **$p < .01$, ***$p < .001$

Despite being a relatively homogeneous group in terms of chronological age, age at onset of deafness was positively correlated with the nonword ratings ($r = +.32$, $p < .01$). Children who lost their hearing at older ages tended to produce nonwords that received higher nonword ratings. This finding suggests that the more opportunity a child has to hear, prior to the onset of deafness, the better their performance will be on tasks involving phonological processing skills such as the nonword repetition task. Duration of deafness was negatively correlated with the perceptual ratings ($r = -.26$, $p < .05$). This finding indicates that children who were deaf for shorter periods of time, prior to implantation, tended to produce nonword utterances that received higher ratings. In addition, we also found that the amount of exposure to oral-only communication was strongly correlated with the nonword perceptual ratings ($r = +.51$, $p < .001$). This relationship demonstrates that oral-aural experiences, promoting the production and perception of spoken language, contribute to phonologically-based tasks such as nonword repetition. The number of active electrodes in the children's implant was not significantly correlated with the nonword perceptual ratings.

Although the nonword repetition task appears on the surface to be simple and straightforward, performance on this task actually reflects the contribution of a number of component subprocesses including initial sensory coding and perception, encoding in working memory, verbal rehearsal, as well as speech articulation and motor control processes in speech production. To examine the contribution of these subprocesses, we also ran correlations between the nonword perceptual ratings and the scores on several standard outcome measures and cognitive processing measures. We expected that the correlations might provide insight into which, if any, skills involved in nonword repetition are indicative of overall performance.

As shown in Table 4, the simple bivariate correlations between the perceptual ratings and several measures of phonological encoding and word recognition (WIPI, LNTe, LNTh, mLNT, and TACL-R) were quite strong and highly significant (r's = +.74, +.77, +.73, +.73, +.50, respectively; p's < .001). In addition, word recognition measured through open-set sentence repetition was related to the nonword ratings. The Bamford-Kowal-Bench (BKB) test was also positively correlated to the perceptual ratings ($r = +.83$, $p < .001$). The children's scores on the VIDSPAC, a measure of speech feature discrimination, were also highly correlated with the perceptual ratings ($r = +.41$, $p < .001$). Similarly, the children's immediate memory span (as measured by the WISC forward digit span task) was significantly correlated with the ratings of their nonword repetitions ($r = +.57$, $p < .001$). The children's performance on the backwards digit span task, however, was not significantly correlated with the perceptual ratings, which is consistent with previous findings that forward and backward digit span tasks measure fundamentally different processes (Lezak, 1983; Li & Lewandowsky, 1995; Rosen & Engle, 1997).

The McGarr intelligibility scores were also strongly correlated with the perceptual ratings of the nonword responses ($r = +.73$, $p <.001$). That is, the intelligibility of children's speech in meaningful sentences was correlated with their performance on the nonword repetition task. In addition, the McGarr durations at all sentence lengths were negatively correlated with the perceptual ratings. The 7-syllable sentences were the most strongly related to the perceptual ratings because they likely provide a better measure of speaking rate than the shorter sentences ($r = -.77$, $p < .001$). Children who spoke more slowly (i.e., had longer sentence durations) on the McGarr task produced poorer imitations of the nonwords.

To control for the known effects of several of the mediating demographic variables, we also ran an additional set of correlations that partialled out the children's age at onset of deafness and duration of deafness. The results of these partial correlations are also shown in Table 4 in the middle column. The similarity of the r-values for the bivariate correlations and the partial correlations indicates that these particular demographic variables had little, if any, effect on the values of the simple bivariate correlations.

Because we found a significant correlation between the children's communication mode scores and their mean perceptual ratings (shown in Table 3), we also ran another set of partial correlations in which the children's age at onset of deafness, duration of deafness, and communication mode scores were partialled out. These results are shown in the far right-hand column of Table 4. The correlation between the perceptual ratings and the word recognition measures, the phonetic feature discrimination measure, the forward digit span, and the speech production measures were all slightly lower than the other two sets of correlations shown in Table 4. This decrease in correlations indicates that the children's mean perceptual ratings were more closely related to their communication mode scores than to their other demographic characteristics. This is further support for the finding that children whose educational environments placed greater emphasis on oral communication received higher mean accuracy ratings for their nonword repetitions.

Although the simple bivariate correlation between the perceptual ratings and the WISC-III backward digit span scores was not significant, these two variables were moderately correlated after we controlled for age at onset of deafness and duration of deafness. This correlation suggests that while age at onset of deafness and duration of deafness are related to the development of phonological working memory, these two variables are not related to the abilities used to perform more complex executive functions indexed by the backward digit span task, such as planning and organizing a problem solving strategy.

Outcome and Process Measures of Performance	Correlation *r*-values Bivariate	Partial 1	Partial 2
Word Recognition			
Word identification, closed-set, pointing response			
Word Intelligibility by Picture Identification (WIPI)	+.74***	+.74***	+.68***
Word identification, open-set, spoken repetition			
Lexical Neighborhood Test, Easy Words (LNTe)	+.77***	+.78***	+.71***
Word identification, open-set, spoken repetition			
Lexical Neighborhood Test, Hard Words (LNTh)	+.73***	+.73***	+.64***
Word identification, open-set, spoken repetition			
Lexical Neighborhood Test, Multisyllabic Words (mLNT)	+.73***	+.72***	+.65***
Sentence identification, open-set repetition of target sentence			
Bamford-Kowal-Bench (BKB)	+.83***	+.84***	+.78***
Language Comprehension			
Closed set, receptive language comprehension of words and sentences			
Test of Auditory Comprehension of Language-Revised (TACL-R)	+.50***	+.51***	+.52***
Phonetic Feature Discrimination			
Perception of speech pattern contrasts			
Video Game for Assessing Speech Pattern Contrasts (VIDSPAC)	+.41***	+.42***	+.31**
Working Memory			
Forward Digit Span			
Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Forward Recall of Digit-Name Lists	+.57***	+.53***	+.46***
Backward Digit Span			
Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Backward Recall of Digit-Name Lists	+.19	+.24*	+.20
Speech Production			
Speech Intelligibility			
McGarr Sentence Intelligibility Test	+.73***	+.74***	+.66***
Speaking Rate			
McGarr Mean Duration of 3-Syllable Sentences (log, msec)	-.54***	-.53***	-.41***
McGarr Mean Duration of 5-Syllable Sentences (log, msec)	-.71***	-.72***	-.65***
McGarr Mean Duration of 7-Syllable Sentences (log, msec)	-.77***	-.77***	-.71***

Table 4. Correlations between nonword repetition perceptual ratings and speech-language measures. Simple bivariate correlations are shown in the left-hand column. Partial correlations controlling for age at onset and duration of deafness are shown in the middle column. Partial correlations controlling for age at onset, duration of deafness, and communication mode are shown in the right-hand column. *$p < .05$, **$p < .01$, ***$p < .001$

Discussion

With regard to the first question that motivated this study, we found a great deal of variability among deaf children with cochlear implants on the nonword repetition task. This variability was revealed by the wide range of perceptual ratings provided by normal-hearing adult listeners. The range of accuracy ratings indicates that the nonword repetition task and perceptual ratings of children's responses were appropriate methods for studying this group of children.

Addressing our second question, we found that individual differences among children in each of the component subprocesses of speech perception, working memory, and speech production were strongly reflected in the nonword repetition scores obtained from the rating procedure. Strong and highly significant correlations were found between the perceptual ratings of the nonword responses and several other measures of word recognition and language comprehension (WIPI, LNTe, LNTh, mLNT, BKB, and TACL-R), working memory (WISC-III forward digit span subtest), speech intelligibility (McGarr Sentence Intelligibility), and speaking rate (McGarr sentence durations). The simple bivariate correlation between the perceptual ratings and the WISC-III backward digit span scores was non-significant, which supports previous findings in the literature suggesting that backward digit span measures different cognitive processes than forward digit span (Li & Lewandowsky, 1995; Lezak, 1983).

In regard to our third question, we found that age at implantation was not significantly correlated with the children's nonword repetition performance. However, two other demographic variables were significantly correlated with the perceptual ratings of the children's imitations, age at onset of deafness and duration of deafness prior to implantation. These two correlations suggest that early exposure to sound and experience hearing and using spoken language are crucial because these activities have subsequent effects on the development of phonological processing skills, language development, and speech production.

In addition to these demographic variables, we found that greater exposure to an exclusively oral educational environment had a significant and beneficial impact on the children's ability to imitate the novel phonological patterns used in the present study. We believe this is a clinically significant finding, because it suggests that oral communication methods may provide deaf children who use cochlear implants with a distinct cognitive processing advantage in learning new vocabulary and developing robust spoken language skills. Passive exposure to speech and spoken language, without explicit analysis and conscious manipulation of the internal structure of these sound patterns, may not be sufficient to develop robust phonological and lexical representations of spoken words and fluency of speech production, especially in children with hearing impairments who routinely receive degraded input signals from their parents and teachers in the language-learning environment.

Phonological Analysis Skills

Deaf children who use cochlear implants may need to be actively engaged in a wide range of activities that involve spoken language processing in order to develop automaticity and automatic attention strategies that can be carried out rapidly, without conscious effort or increased processing resources. This may be one important and direct benefit of oral-only education programs. The excellent spoken language skills acquired by children in oral-only programs may reflect development of explicit phonological analysis skills, permitting the child to engage in active perceptual processing strategies. These phonological analysis skills include processing activities such as "decomposition" of a speech pattern into a sequence of discrete phonological units and then "reassembly" of those individual units into sequences of highly coordinated gestures that serve as input to motor control processes used in speech production.

Explicit phonological coding skills of this kind may result in increases in the speed and efficiency of constructing phonological representations of spoken words in short-term memory. Recovering and maintaining the structural description of a novel phonological pattern in speech perception and then reconstructing the same phonological pattern again in speech production may be fundamental to establishing permanent links between speech perception and production. These complementary links may then lead to further development of highly automatized sensory-motor articulatory programs for covert verbal rehearsal and phonological coding in working memory. Thus, the superior phonological processing skills of oral deaf children using OC may simply be a natural byproduct of the extensive emphasis on speech and spoken language skills in oral-only educational environments. The phonological analysis skills gained from experience and activities in oral-only environments may account for why oral-only children consistently display better scores on a wide range of outcome measures of speech and language than children from TC environments.

The results of this study revealed relatively large and consistent effects of early experience on nonword repetition performance in deaf children who use CIs. But exactly how does the early sensory and linguistic experience in the language learning environment affect a deaf child's information processing skills after receiving a cochlear implant? This is a fundamental research question that clearly touches on much broader topics such as learning, memory, and cognition. Although it has been well documented over the years that a deaf child's communication mode plays an important role in the development of speech and oral language skills following implantation (Kirk, Pisoni, & Miyamoto, 2000; Osberger, Robbins, Todd, & Riley, 1994; Svirsky, Robbins, Kirk, Pisoni, & Miyamoto, 2000; Tait, Lutman, & Robinson, 2000), there has been little discussion of the underlying neural and cognitive mechanisms responsible for these differences in performance.

Oral vs. TC Environments

Numerous studies have reported that deaf children immersed in oral-only programs do much better on a range of oral language tasks than deaf children who are in TC programs (Kirk et al., 2000; Osberger et al., 1994; Svirsky et al., 2000). Such studies show that oral-only children perform better than TC children on tests assessing speech feature discrimination, comprehension, spoken word recognition, receptive and expressive language, and speech intelligibility. Why does this difference occur? What factors are responsible for the better performance of oral-only children? Can we provide a theoretically motivated explanation of these differences that is consistent with what we know about speech and language development?

As described earlier, among other findings, we found that the nonword repetition scores were strongly correlated with digit span scores. However, this result was highly selective in nature and was only observed for the forward digit span scores and not the backward scores. This pattern of results suggested that early oral experience in the language learning environment may influence phonological coding and rehearsal processes in short-term memory but may not affect controlled attention and executive functions that are used for active processing in working memory (Lezak, 1983; Li & Lewandowsky, 1995; Rosen & Engle, 1997).

We also found that the nonword repetition scores were negatively correlated with speaking rate, which can be taken as an index of the speed of covert verbal rehearsal that is used to maintain phonological representations in working memory. Children from oral-only environments not only received higher nonword repetition ratings in this study, but in previous studies, they also displayed longer forward digit spans, and faster speaking rates than the children from TC environments (Burkholder, this volume; Pisoni & Cleary, in press). These results suggest a common locus in the information processing system for the effects of early experience on speech and language development. This locus appears to be associated with the verbal rehearsal mechanism that is used to maintain phonological representations of sound patterns in short-term memory.

Why do children from TC environments consistently perform worse on a wide range of speech and language outcome measures than the children in Oral environments? Several reasons can be offered to explain the differences in performance between the two groups of children. These factors may provide some insights into how the nature of the child's early sensory experience and activities affects the speed and efficiency of the underlying processing mechanisms that are used in a range of behavioral tasks. First, children in TC environments routinely use simultaneous communication. For these children, speech is combined with some form of manual communication such as Signed Exact English (SEE). One consequence of using simultaneous communication is that the rate of articulation and verbal rehearsal and the transfer of linguistic information are much slower than using speech alone without any manual signs (Bellugi & Klima, 1997; Emmorey, 2002).

A second reason for the poorer performance of TC children may be due to competition between speech and manual communication for controlled attention and limited processing resources of short-term memory. Under simultaneous communication, speech and sign do not specify the same gestures and common underlying articulatory events. The information from the two input modalities is not congruent as it is when a listener sees the talker's lips and hears his speech under auditory-visual presentation conditions. Moreover, because the child is looking at the talker's hands, lip reading is not the primary means of providing complementary phonetic information about the speech signal. Thus, little, if any, additional facilitation is gained from the visual input; if anything there are strong reasons to support the argument that competition and inhibition result from the presentation of two divergent input signals (see Doherty-Sneddon, Bonner, & Bruce, 2001; Tyler, 1993).

These observations about multimodal speech perception are consistent with several recent findings showing interference and inhibition effects of manual communication skills in TC children who are learning oral language via their cochlear implant (Bergeson, Pisoni, & Davis, in progress). In TC children, knowledge and use of manual language competes with the dominant mode of processing speech, via the auditory sensory modality. Differences in input modality between sign language and speech make it difficult for TC children to integrate common gestural information across the two sensory modalities. This information "mismatch" increases the processing load on the working memory system that plays an important role in language comprehension and word recognition (Baddeley, Gathercole, & Papagno, 1998). Thus, it is not surprising that TC children generally show lower scores on a variety of speech and language assessments that measure auditory-only or auditory and visual language processing skills.

Third, it is possible that deaf children who receive cochlear implants have hearing parents who often do not have an expert knowledge of signs and manual communication. This is a reasonable assertion given that about 90 percent of deaf signers without cochlear implants are estimated to have had exposure to atypical signing (Woll et al., 2002). One consequence of atypical signing exposure is that the language model that deaf children in a TC environment experience is likely to be incomplete or impoverished (see Kluwin, 1981). This means that some TC children may not get exposed to robust experiences of manual sign and may have difficulties decoding the morphology and syntax of the language.

Finally, it is also possible that the differences in early speech and language experience after implantation between the TC and oral-only children may produce several different effects on the initial encoding of speech signals and subsequent rehearsal processes used to maintain phonological representations of spoken words in short-term memory. Because TC children have less exposure to speech and spoken language in their early environment after implantation, they may actually have two separate but closely related problems in processing and rehearsing auditory information in short-term memory. The first problem concerns initial encoding and recognition. The lack of exposure to speech may affect the development of automatic attention and the speed with which speech signals can be rapidly identified and recoded into phonological representations in short-term memory. As a consequence, TC children may also have difficulty in scanning and retrieving phonological representations of spoken words in short-term memory.

The second problem deals with the verbal rehearsal process that maintains phonological information in working memory. TC children may not only have slower scanning processes but they may also have slower and less efficient verbal rehearsal strategies once phonological information finally gets into short-term memory simply because they have had less experience in producing speech and spoken language. Encoding and retrieval processes that are very fast and highly automatized in normal-hearing children may be much slower and more effortful for deaf children with CIs, particularly those deaf children who use TC methods.

Scanning of Short-Term Memory

The proposal that deaf children with CIs, especially deaf children who use TC, have atypical abilities to verbally rehearse and scan phonological information in short-term memory was recently confirmed in our laboratory by Burkholder (this volume). In a study using McGarr sentence durations and speech timing measures made during WISC-III digit span recall, she found that children with cochlear implants had slower speaking rates and shorter digit spans than their NH peers. Within the group of children with cochlear implants, she found that TC children had much slower speaking rates than OC children but they also had shorter digit spans. Both findings support the hypothesis that the TC children have slower subvocal verbal rehearsal rates, which may contribute to their shorter digit spans and impact other speech and language processing tasks as well.

In addition, Burkholder (this volume) found that the CI children displayed longer interword pauses during digit span recall than NH children. Cowan (1992; 1999) has suggested that interword pause times during immediate recall can be used as a measure of scanning or retrieval speed. Therefore, longer interword pauses indicate a slower speed of scanning or retrieving information. Within the CI group, the TC users appear to be scanning items in short-term memory more slowly, because the interword pauses displayed in this group were longer than in the OC group. Thus, in addition to slower verbal rehearsal rates, slower scanning and retrieval speeds are also likely to be responsible for the shorter digit spans observed in the TC users and the poorer performance in a range of other speech and language tasks.

Future Directions

Future research should also consider more closely how verbal rehearsal and scanning of short-term memory operate in the CI population in a variety of other phonological working memory tasks. It would be useful to expand the current study to include speech-timing measures in the CI children's nonword repetitions. Examining speaking rate would aid in determining if the scanning processes involved in completing nonword repetition are also slower in TC children. Based on the present findings that TC children performed worse on nonword repetition than OC children, it is likely that TC children also have slower scanning and retrieval processes. These slower scanning speeds, along with slower subvocal rehearsal speeds, could be related to the TC children's lower nonword accuracy ratings, in addition to the demographic and speech perception variables that were found to be correlated to nonword repetition accuracy ratings in the current study.

Another area for future investigation is related to the finding that the children in this study imitated the shorter nonwords more accurately than the longer nonwords. In light of this finding and previous results showing that the phonological structure of the nonword pattern is related to the children's ability to accurately imitate it (Cleary et al., 2002; Carter et al., in press), future studies of repetition performance involving nonword patterns balanced in terms of stress patterns and phonemic content could provide further insights into the nature of the underlying phonological systems of deaf children after cochlear implantation and how these systems may be different from those of normal-hearing children.

All of these questions would be interesting to study in normal-hearing children as well, using a variety of new experimental techniques. In future work, we are planning to study normal-hearing children in quiet conditions and under conditions using processing strategies designed to simulate the degraded speech signals provided to deaf children by their cochlear implants (Dorman, Loizou, Kemp, & Kirk, 2000; Kaiser & Svirsky, 2002). Simulation studies like these using normal-hearing typically developing children should provide new insights into how the nonword repetition task is carried out. In addition, comparison of the normal-hearing children's performance in simulation studies to the performance of deaf children who use CIs should provide valuable new knowledge about how the absence of sound during early stages of speech and language development affects the speed and efficiency of the phonological processing skills used to rapidly encode, maintain and reproduce novel phonological patterns from working memory.

Conclusions

The results of the present study indicate that deaf children who use cochlear implants demonstrate wide variability in spoken language skills used to imitate novel nonword patterns. The observed differences in performance can be accounted for, in part, by variation in the separate underlying cognitive processes involved in spoken word recognition, comprehension, working memory, and verbal rehearsal speed as indexed by speaking rate. In addition, the efficiency and speed of these cognitive processes seems to be related to the nature of the early auditory and linguistic experiences and activities that the CI users are engaged in after receiving their CIs. This study revealed that deaf children who use oral communication (OC) methods received higher nonword accuracy ratings than deaf children who use simultaneous speech and signing methods (TC). We suggest that the increased amounts of early linguistic experience and processing activities provided to OC users is responsible for their ability to rapidly encode, represent, subvocally rehearse, retrieve, and produce the novel phonological patterns presented to them in a nonword repetition task.

References

Baddeley, A.D., Gathercole, S.E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Bellugi, U., & Klima, E.S. (1997). Language, spatial cognition, and the brain. In M. Ito, Y. Miyashita, & E. Rolls. *Cognition, computation, and consciousness* (pp. 177-189). London: Oxford University Press.

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.

Bergeson, T.R., Pisoni, D.B., & Davis, R.A.O. (in progress). A longitudinal study of audiovisual speech perception by prelingually deaf children with cochlear implants.

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. 251-275). 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*.

Cleary, M., Dillon, C.M., & Pisoni, D.B. (2002). Imitation of nonwords by deaf children after cochlear implantation: Preliminary findings. *Annals of Otology, Rhinology & Laryngology, 111*, 91-96.

CoolEdit version 96. (1996). Syntrillium Software Corporation. Phoenix, AZ.

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language, 31*, 668-684.

Cowan, N. (1999). The differential maturation of two processing rates related to digit span. *Journal of Experimental Child Psychology, 72*, 193-209.

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.

Doherty-Sneddon, G., Bonner, L., & Bruce, V. (2001). Cognitive demands of face monitoring: Evidence for visuospatial overload. *Memory & Cognition, 29,* 909-919.

Dorman, M.F., Loizou, P.C., Kemp, L.L., & Kirk, K.I. (2000). Word recognition by children listening to speech processed into a small number of channels: Data from normal hearing children and children with cochlear implants. *Ear & Hearing, 21*, 590-596.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research.* Mahwah, NJ: Lawrence Erlbaum Associates.

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

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 & Cognition, 23*, 83-94.

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., Spehar, B., Torretta, G., Tobey, E., 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). St. Louis, MO: Central Institute for the Deaf.

Kaiser, A.R., & Svirsky, M.A. (2002). Using a personal computer to perform real-time signal processing in cochlear implant research. Unpublished Manuscript.

Kirk, K.I., Pisoni, D.B., & Miyamoto, R.T. (2000). Lexical discrimination by children with cochlear implants: Effects of age at implantation and communication mode. In S. Waltzman, & N. Cohen (Eds.), *Cochlear Implants* (pp. 252-254). New York: Thieme.

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.

Kluwin, T.N. (1981). The grammaticality of manual representations of English in classroom settings. *American Annals of the Deaf, 126*, 417-421.

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.

McGarr, N. (1981). The effect of context on the intelligibility of hearing and deaf children's speech. *Language and Speech, 24*, 255-263.

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

Pisoni, D.B., Cleary, M., Geers, A.E., & Tobey, E.A. (2000). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *The Volta Review, 101*, 111-164.

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 (Suppl.), 18*, 592-593.

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.

Svirsky, M.A., Robbins, A.M., Kirk, K.I., Pisoni, D.B. & Miyamoto, R.T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science, 11,* 153-158.

Tait, M., Lutman, M.E. & Robinson, K. (2000). Preimplant measures of preverbal communicative behavior as predictors of cochlear implant outcomes in children. *Ear & Hearing, 21*, 18-24.

Tyler, R.S. (1993). Speech perception by children. In R. Tyler (Ed.), *Cochlear Implants* (pp. 191-256). San Diego: Singular.

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

Woll, B., Grove, N., Atkinson, J., Morgan, G., Smith, N., Tsimpli, I., & Sieratzki, H. (2002). *Exploring issues of language disorders and modality through the study of atypical signers*. Symposium conducted at the meeting of the 23rd Symposium on Research in Child Language Disorders and 9th Congress International Association for the Study of Child Language, Madison, Wisconsin. July 16-21, 2002.

Measures of Working Memory Span and Verbal Rehearsal Speed in Deaf Children Following Cochlear Implantation[1]

David B. Pisoni and Miranda Cleary

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

[1] This work was supported by NIH-NIDCD Training Grant T32DC00012 and NIH-NIDCD Research Grants R01DC00111, NIH NIDCD R01DC00064 to Indiana University and NIH Research Grant R01DC03100 to Central Institute for the Deaf. We would like to thank Ann Geers, Chris Brenner and Mike Strube of CID for their help and assistance. We also thank Tonya Bergeson, Steven Chin and Karen I. Kirk for comments on an earlier version of this paper. Several portions of this report are based on a longer chapter describing related research to be published in Springer Handbook of Auditory Research: Auditory Prostheses, SHAR Volume X. F.G. Zeng, A.N. Popper & R.R. Fay (Eds.).

Measures of Working Memory Span and Verbal Rehearsal Speed in Deaf Children Following Cochlear Implantation

Abstract. Large individual differences in spoken word recognition performance have been found in deaf children following cochlear implantation. Recently, Pisoni and Geers (2000) reported that simple forward digit span measures of verbal working memory were significantly correlated with spoken word recognition scores even after potentially confounding variables were statistically controlled for. The present study replicates and extends these initial findings to the full set of 176 participants in the CID cochlear implant study. The pooled data indicate that despite statistical "partialling-out" of differences in chronological age, communication mode, duration of deafness, duration of device use, age of onset of deafness, number of active electrodes, and speech feature discrimination, significant correlations still remain between digit span and several measures of spoken word recognition. Strong correlations were also observed between speaking rate and both forward and backwards digit span, a result that is similar to previously reported findings in normal-hearing adults and children. The results suggest that perhaps as much as 20% of the currently unexplained variance in spoken word recognition scores may be independently accounted for by individual differences in cognitive factors related to the speed and efficiency with which phonological and lexical representations of spoken words are maintained in and retrieved from working memory. A smaller percentage, perhaps about 7% of the currently unexplained variance in spoken word recognition scores, may be accounted for in terms of working memory capacity. We discuss how these relationships may arise and their contribution to subsequent speech and language development in prelingually deaf children who use cochlear implants.

Introduction: Individual Differences and Variation in Outcome

Despite the success of cochlear implants in many prelingually deafened, early-implanted children, enormous individual differences have been reported on a wide range of speech and language outcome measures. Some children do extremely well with their cochlear implant while others derive only minimal benefits. Although large individual differences in outcome following implantation have been well documented for many years in the clinical literature, the factors responsible for variation in performance are still not well understood (Blamey et al., 2001; Hodges et al., 1999; Kirk, 2000; Pisoni, 2000; Sarant, Blamey, Dowell, Clark, & Gibson, 2001). Identifying the reasons for the wide variability in outcome measures following cochlear implantation is a challenging research problem because a large number of complex sensory, perceptual, cognitive and linguistic processes affect speech and language performance in any particular behavioral task. It may be fruitful to investigate these complex interactions directly using measures that assess individual component processes of speech and language behavior, if we want to explain why some pediatric cochlear implant users do so well while others struggle and achieve only small benefits after receiving a cochlear implant.

The observed individual differences can be extremely striking. In some of our earlier research we have looked in detail at several of the factors distinguish children who display exceptionally good performance with their cochlear implants, from those that derive only minimal benefit (Pisoni, Cleary, Geers & Tobey, 2000). The children who show exceptional progress appear to acquire spoken language quickly and easily and seem to be on a developmental trajectory that parallels children with normal hearing (Svirsky, Robbins, Kirk, Pisoni, & Miyamoto, 2000). The exceptionally good performance of these so-called "Stars" is not merely an anomaly, but rather can be viewed as an "existence proof" for the best-case scenario offered by cochlear implantation of prelingually deafened children: cochlear implants work very well for some pediatric cochlear implant users, greatly facilitating the processes of speech perception and language development. The problem is that cochlear implants do not work well with all deaf children. Why does this occur? What factors are responsible for the enormous variation in outcome? These are two of the primary questions we have pursued in our research program on individual differences.

Among implanted children in general, approximately 40-65% of the existing variance in outcome measures can be accounted for in terms of a small number of traditional demographic variables such as duration of deafness, length of device use, age at implantation, and residual hearing prior to implantation (Blamey et al., 2001; Dowell, Blamey, & Clark, 1995; Miyamoto et al., 1994; Sarant et al., 2001; Snik, Vermeulen, Geelen, Brokx, & van den Broek, 1997; Zwolan et al., 1997). Although the role played by these familiar demographic variables has been extensively studied over the last few years, it has become clear from the substantial amount of remaining unexplained variability, that further research using a new theoretical framework is needed to understand the large individual differences in performance outcome. Our earlier work has shown that the performance of the "Stars" cannot be explained merely by appeal to demographics alone: children with similar demographic backgrounds and medical histories often show widely differing degrees of success with an implant (Balkany, Hodges, Miyamoto, Gibbin, & Odabasi, 2001; Geers et al., 1999; Kirk, 2000; Pisoni, 2000; Pisoni, Svirsky, Kirk, & Miyamoto, 1997). New techniques and methodologies are needed at this time to reveal the source of these individual differences in speech and language outcomes.

One path of investigation has offered some promising new insights into the nature of the observed individual differences. The examination of "higher-level" perceptual, cognitive and linguistic factors has typically not received much attention in the field of cochlear implants until recently (Pisoni, 2000). One reason for the lack of knowledge about central auditory factors and cognitive processes in this clinical population is that most of the research on cochlear implants has been carried out by medical professionals who have been concerned primarily with questions of device "efficacy" and assessment of outcome and benefit following implantation. Research on device efficacy requires well-defined assessment measures of outcome performance that are familiar to surgeons and clinicians who work with deaf patients. However, in order to specify how various central factors may combine to determine the degree of benefit observed in individual patients following cochlear implantation, it is necessary to go beyond basic clinical assessment using only traditional audiological tools.

Over the last few years, we have begun to focus on questions surrounding variability in outcome using several new methodologies that go beyond the standard end-point assessment-based clinical measures usually administered following cochlear implantation. The results of our research suggest that additional sources of variance may indeed come from more central cognitive factors (Pisoni et al. 2000; Pisoni et al., 1997). An enormous amount of information processing takes place beyond the auditory periphery for spoken language understanding to occur, and we suggest that the time is ripe to focus on the process of how deaf children with cochlear implants are able to use the initial sensory input conveyed by their devices. That is, research efforts should examine not only what the children "hear" but what they are able to do with the sensory information provided by their cochlear implants.

For spoken language processing to proceed rapidly and efficiently, it is not only essential that auditory information be made available to the central auditory system, but also that once encoded, this information can then be reliably maintained, retrieved, and transformed into phonological and lexical representations for use in a range of different speech and language processing tasks. These cognitive processing abilities are not sprung fully formed in the human infant, but rather, even in normal-hearing children, develop over time as a result of experience-dependent learning (see Jusczyk & Luce, 2002; Locke, 1993). To understand the development of these speech and language processes in deaf children with cochlear implants, it is important to understand how the language-learning environment modulates, shapes, and facilitates the developmental process. It is probably reasonable to suppose that some of the most radical neural changes that take place after cochlear implantation in order to make speech perception possible, occur at quite high levels of central auditory processing--not at the level of auditory periphery. Thus, our research on cochlear implants in deaf children has focused on the underlying basic cognitive information processing skills that are used to support the perception and production of spoken language (see Pisoni, 2000).

The Information Processing Approach to Cognition

In order to gain a better understanding of what deaf children are learning after they receive a cochlear implant and how they use auditory sensory input, we have adopted an information processing approach which looks closely at the content and flow of information within the nervous system and how it changes over time (Pisoni, 2000). The foundational assumption of this approach is to view the human nervous system as an information processor. An information processor is a system that encodes, stores and manipulates various types of symbolic representations (Haber, 1969; Lachman, Lachman, & Butterfield, 1979). Information exists in several different forms at a number of levels of representation in the nervous system ranging from early registration and encoding of the sensory input to permanent storage of neural and linguistic representations in long-term memory.

By viewing the mechanisms of sensation, perception, attention, memory and learning within this larger integrated framework of information processing, cognitive scientists have gained many new insights into the structure and function of the central nervous system (Neisser, 1968; Reitman, 1965). The information processing framework has also stimulated the development of new tools and experimental methodologies to study the processes that underlie these behaviors and has led to new theoretical conceptualizations that can be used to explain and predict variation and variability in more complex higher-level behaviors such as speech and language in different populations. This approach to human cognition has also provided researchers with the theoretical framework for reformulating some long-standing problems as well as identifying new research questions. The results of these efforts have provided fundamental new knowledge about perceptual and cognitive development and the neural processing mechanisms that underlie behavior (Gazzaniga, 2000; Posner, 1989).

One of the most important and influential proposals that has emerged from the information processing framework is the construct of working memory. Working memory is usually defined as a mechanism for the temporary storage and early processing of sensory and perceptual information (Baddeley, 1992; Baddeley & Hitch, 1974). Although this system is markedly limited in its capacity and can be subject to rapid forgetting and loss, it serves a vital role in temporarily maintaining information for further processing. Because it is posited that information must be represented in working memory before it can be more permanently stored in long term memory, working memory has been described as serving as the "interface" or "gateway" between the initial sensory input and stored knowledge, i.e., learned material.[2]

One traditional method of assessing individual differences in working memory capacity is to find the number of familiar items that can be recalled in correct serial order. Digit span is the most widely used measure of verbal working memory capacity, and is often administered using two different variants:

[2] Although we are here using the terms "working memory," "short-term memory" and "immediate memory" interchangeably, the term "working memory" is sometimes reserved for tasks that require the maintenance of information while additional new information is presented for processing. By this latter view, verbal digit span is better described as a simple verbal short-term or immediate memory task.

"forward" digit span requires simple verbatim recall of the list of digits to be remembered, while "backwards" digit span requires the subject to reproduce a given target list with the items in reverse order (Wechsler, 1991).

Although the temporary nature of the working memory system and its limited capacity render these initial representations quite fragile, there are ways of circumventing these limitations. Rehearsal is a generic term used in cognitive psychology to refer to methods for maintaining information in working memory via "refreshing" or re-encoding of the material to-be-remembered (Atkinson & Shiffrin, 1968). One ubiquitous rehearsal method for normal-hearing adults is verbal rehearsal--simple vocal or subvocal (internal/silent) repetition of the verbal materials to be remembered. Interestingly, it has been found that although normally-functioning adults typically rehearse "silently" or sub-vocally "to themselves," measures of actual articulation speed tend to correlate well with the rates at which this internal verbal rehearsal is carried out (Landauer, 1962; Standing & Curtis, 1989). Verbal rehearsal and its relationship to overt articulation appears gradually in normal development, and begins to be employed by normal-hearing children between the ages of five and seven years (Flavell, Beach, & Chinsky, 1966; McGilly & Siegler, 1989).

A number of recent findings have suggested that investigating the properties of working memory may provide new insights into the nature and locus of the individual differences observed among users of cochlear implants (see Carpenter, Miyake & Just 1994; Gupta & MacWhinney 1997; Baddeley, Gathercole, & Papagno 1998). In this report, we extend and expand upon some of the preliminary findings on working memory reported in Pisoni and Geers (2000) and Cleary, Pisoni, Kirk, Geers, and Tobey (2000).

Pisoni and Geers (2000) found that among a group of 43 pediatric cochlear implant users with a relatively homogeneous demographic background, working memory measures of verbal digit span showed strong positive correlations with measures of speech perception, speech production, language development, and reading skills. In follow-up analyses, Pisoni, Cleary, Geers, and Tobey (2000) reported that individual differences in verbal digit span were strongly correlated with a measure of articulation time obtained from a separate speech production task. (The children reported on in these previous papers are a subset of the current group of children.)

The results of these two earlier studies suggested that it might be informative to examine how processing capacity and verbal rehearsal speed each contribute to the relationship found between verbal digit span and several of the outcome measures of interest. For the purposes of the present paper, we focus primarily on the relationship between digit span and spoken word recognition. Our focus on spoken word recognition (as opposed to phoneme discrimination or auditory sentence comprehension) is motivated by the tendency for performance on this task to effectively separate out children who are receiving minimal benefit from their implant from those who are successfully acquiring spoken language (Pisoni, Svirsky, Kirk, & Miyamoto, 1997).

Methods and Procedures

To obtain measures of working memory capacity from a large number of deaf children following cochlear implantation, we were fortunate to be able to collaborate with Ann Geers and her colleagues at Central Institute for the Deaf (CID) in St. Louis who already had an on-going large-scale research project underway (Geers & Brenner, in press; Geers, Brenner, & Davidson, in press). The CID project was designed to obtain a wide range of outcome measures of speech, language and reading skills from 8- and 9-year old children who had all used their cochlear implants for at least three and one-half years. Thus, chronological age and length of implant use were relatively controlled within the sample of children studied.

Using the test lists and procedures from the WISC III (Wechsler 1991), forward and backward auditory digit spans were obtained from four groups of 8- and 9-year-old children with cochlear implants. A total of 176 pediatric cochlear implant users were individually tested in separate groups at CID in St. Louis during the summers of 1997, 1998, 1999 and 2000. Forward and backward digit spans were also collected from an additional group of 45 age- and gender-matched normal-hearing 8- and 9-year old children who were tested in Bloomington, Indiana and served as a comparison group.

The WISC-III digit span memory task requires the child to repeat back a list of digits that are spoken live-voice by an experimenter at a rate of approximately one digit per second (WISC-III Manual, Wechsler 1991). In this study, the digit span task was administered with the face of the clinician visible to the child. For the "digits-forward" section of the task, the child was required to simply repeat back the list as heard. For the "digits-backward" section of the task, the child was told to "say the list backward." In both parts of the procedure, the lists began with two items, and were increased in length upon successful repetition until a child got two lists incorrect at a given length, at which point the testing stopped. Points were awarded for each list correctly repeated with no partial credit. Each child's digit span in points was calculated by summing the number of lists correctly recalled at each list length. The total points score for forward digit span could vary between zero and 16; the total points score for backward digit span could vary between zero and 14.

Results

Figure 1 shows the frequency distributions for the forward and backward spans for all 176 pediatric cochlear implant users. The top panel shows the forward spans; the bottom panel shows the backwards spans. The distributions shown in Figure 1 closely approximate normal distributions and provide reassurance that the difficulty of the task was appropriate for the children tested.

A summary of the mean digit span results for all five groups of children is shown in Figure 2. Forward and backward digit spans are shown separately for each group. The children with cochlear implants are shown in the four panels on the left by year of testing; the mean scores for the group of normal-hearing children are shown on the right.

Inspection of the data shown in Figure 2 reveals an orderly and systematic pattern of the forward and backward digit spans for the deaf children with cochlear implants. All four groups are quite similar to each other. Within each group, the mean forward digit span is clearly longer than the mean backward digit span. The pattern is quite stable over the four years of testing despite the fact that these scores are based on independent groups of subjects. The difference in span length between forward and backward report was highly significant for the entire group of 176 deaf children and for each group taken separately ($p < .001$).

The mean forward and backward digit spans obtained from the group of 44 age-matched normal-hearing children are shown in the right-hand panel of Figure 2. Examination of these data show that the digit spans for the normal-hearing children differ in several ways from the digit spans obtained from the children with cochlear implants. Firstly, although the mean digit spans for the normal-hearing children shown of Figure 2 are age-appropriate based on the published norms for the WISC III (Wechsler, 1991), the mean spans for the cochlear implant group are noticeably lower than would be expected from the published norms. For both forward and backward digit span, the normal-hearing children display longer spans than those obtained from the deaf children with cochlear implants. The difference is especially marked in the case of forward digit spans. This finding suggests possible atypical development of short-term memory capacity in the children with cochlear implants and indicates possible differences between the two groups in the underlying processing mechanisms that are used to encode and maintain sequences of spoken digits in immediate memory.

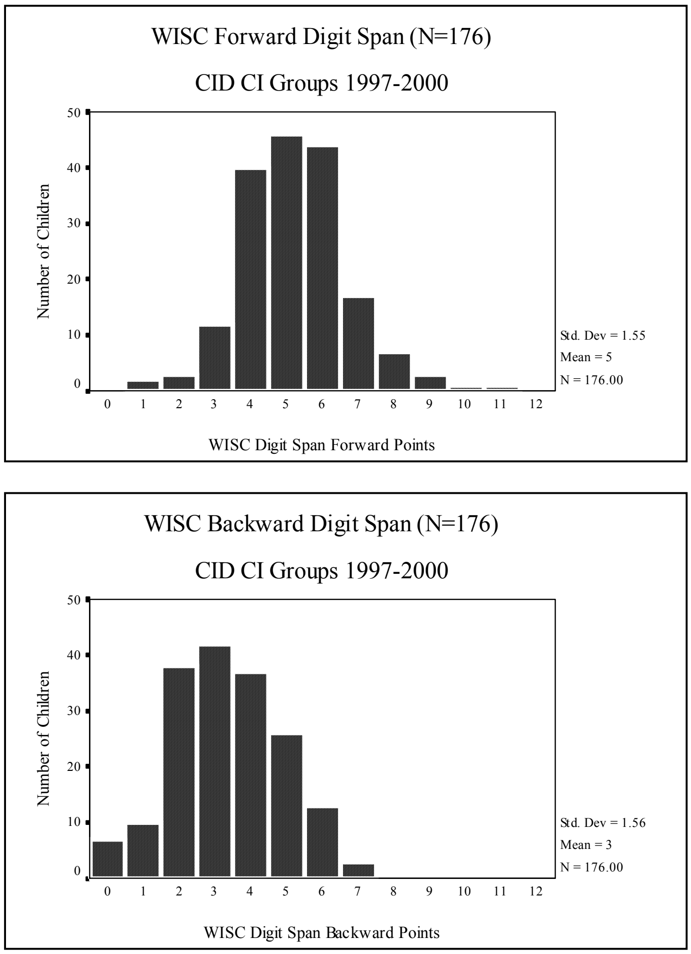


Figure 1. Frequency histograms of WISC digit span scored by points for the 176 8- and 9-year old children with cochlear implants. Forward digit spans are in the top panel, backwards digit spans in the bottom panel.

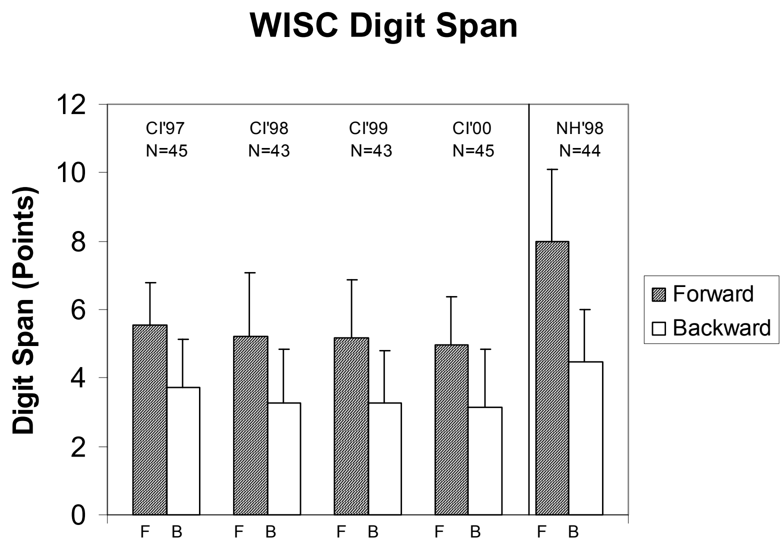


Figure 2. Mean WISC digit spans scored by points for the four groups of 8- and 9-year old children with cochlear implants and for a comparison group of 8- and 9-year-old normal-hearing children. Forward digit spans are shown by the shaded bars, backwards digit spans by the open bars. Error bars indicate one standard deviation from the mean.

Digit Span and Communication Mode

In order to account for the observed differences in auditory digit span among the children with cochlear implants, we examined the correlations between digit span and several of the traditional demographic variables such as age of onset of deafness, duration of deafness, age at implantation, duration of implant use, communication mode, age, gender, and number of active electrodes. Of the various demographic measures available, the only one that correlated notably with digit span was communication mode. Communication mode refers to the nature of the child's early sensory and linguistic experience after receiving a cochlear implant and is here indexed by the degree of emphasis on oral versus manual language skills by parents, teachers and therapists in the home and in the child's educational environment (Geers et al. 1999; Geers & Brenner, in press).

To determine communication mode, each child's degree of exposure to Oral-only communication methods was quantified by determining the type of communication environment experienced by the child in the year just prior to implantation, each year over the first three years of CI use, and then in the year just prior to the current testing. A score was then assigned to each year, with a "1" corresponding to the use of "total communication" with a sign emphasis (that is, extensive use of manual signs in addition to spoken language), and a "6" indicating an auditory-verbal environment with a strong emphasis on auditory communication without the aid of lipreading (see Geers et al. 1999; Geers & Brenner, in press for details). Communication methods intermediate between these two extremes were assigned intermediate scores ranging from 2 to 5. These scores were then averaged over the five points in time. The mean communication mode score for the group over the five intervals was approximately 3.9 on this 6-point scale. A wide range of communication mode backgrounds was however present within the sample (range of average communication mode scores = 1.0 to 6.0).

We found that forward digit span was positively correlated with Communication Mode ($r = +.34$, $p < .001$). Children who were in language learning environments that primarily emphasized oral skills tended to display longer forward digit spans than children who were in total communication (TC) environments. However, the correlation between digit span and communication mode was selective in nature because its statistical significance was restricted only to the forward digit span scores; the backward digit spans were not significantly correlated with communication mode ($r = +.14$, $p = .06$) or with any other demographic variable, except for chronological age at time of testing ($r = +.22$, $p < .01$).

In order to further examine the effects of early experience on working memory spans, a median split was carried out on the communication mode scores to create two subgroups, Oral children and TC children. Figure 3 shows the digit spans plotted separately for the Oral and TC children for each of the four years of testing at CID. Examination of the forward and backward digit spans for these two groups of children indicates that the Oral groups consistently displayed longer average forward digit spans than the TC groups. While the differences in mean forward digit span between Oral and TC groups were highly significant ($p < .001$), the differences in backward digit span were not (e.g., for $N = 176$, $p = .22$).

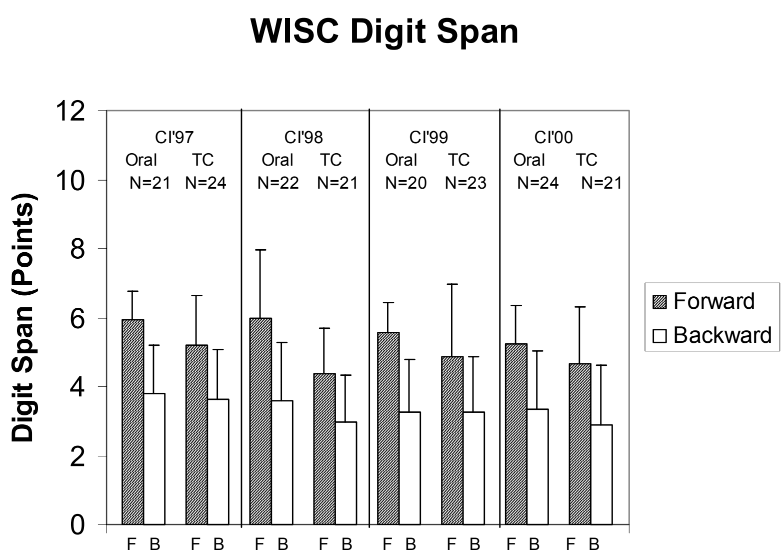


Figure 3. Mean WISC digit spans scored by points for the four groups of 8- and 9-year old children with cochlear implants, separated by communication mode. For each year, scores for the oral group are shown to the left of those for the total communication group. Forward digit spans are shown by the shaded bars, backwards digit spans by the open bars. Error bars indicate one standard deviation from the mean.

The difference in forward digit span between Oral and TC children is present at each year of testing and suggests that forward digit spans are sensitive to the nature of the early sensory and linguistic experience that the child receives immediately after cochlear implantation. The differences observed in the forward digit spans could be due to several factors, such as better encoding of the initial stimulus patterns into more stable and robust phonological and lexical representations in working memory, greater speed and efficiency of the verbal rehearsal processes that are used to maintain information in working memory, or possibly even faster rates of retrieval of information from working memory during recall. All three factors could influence measures of information processing capacity and any one of these could affect the number of digits correctly recalled from immediate memory in this task.

Regardless of which factor or factors are responsible for the differences, the present results demonstrate that forward digit span is sensitive to the effects of early experience and suggest that several specific mechanisms in the information processing system may be affected by the nature of the early experience the child receives after implantation. Although these results indicate that early experience in an environment that emphasizes oral language skills is associated with increased information processing capacity in verbal working memory, additional converging measures of performance would be helpful to specify precisely what elementary processes and information processing mechanisms are responsible for the longer forward digit spans observed in these children.

Digit Span and Spoken Word Recognition

While traditional demographic factors such as duration of deafness, length of device use and age at implantation have been shown to be related to individual differences in spoken word recognition, we suggest that a portion of the remaining unexplained variance can be accounted for in terms of individual differences in information processing capacity as measured by verbal digit span. Numerous studies of normal-hearing children over the past few years have demonstrated close "links" between verbal short-term/working memory and learning to recognize and understand new words (Gupta & MacWhinney 1997; Gathercole, Hitch, Service, & Martin, 1997). More specifically, it has been demonstrated that individual differences in the ability to imitate sound forms of novel pseudo-words are positively correlated with individual differences in vocabulary and novel word learning (Baddeley, Gathercole, & Papagno 1998; Gathercole, Hitch, Service, & Martin, 1997). Other research (e.g., Adams & Gathercole, 2000) has suggested that important milestones in speech and language acquisition are associated with developmental changes in verbal working memory.

To determine if measures of working memory capacity are related to spoken word recognition in deaf children following cochlear implantation, we correlated the WISC forward and backward digit span scores with three different measures of spoken word recognition performance that were obtained from the children tested at CID. A summary of the correlations between digit span and word recognition scores based on these 176 children is shown in Table I for the WIPI, LNT and BKB word recognition tests.

The WIPI (Word Intelligibility by Picture Identification Test) is a closed-set test of auditory word recognition in which the child selects a word from among six alternative pictures (Ross & Lerman, 1979). The LNT is an open-set test of word recognition and lexical discrimination that requires the child to imitate and reproduce an isolated word (Kirk, Pisoni & Osberger, 1995). This test is similar to the well-known PBK test although the vocabulary on the LNT was designed to control for familiarity while lexical competition among the items was manipulated systematically to measure discrimination among phonetically similar words in the child's lexicon. Finally, the BKB is an open-set word recognition test in which key words are presented in sentence contexts (Bench, Kowal & Bamford, 1979). For the CID study, all of the word recognition test materials were pre-recorded and presented in the auditory-only modality.

Table I displays two sets of correlations. The left-hand portion of the table shows the simple bivariate correlations of the forward and backward digit spans with the three measures of spoken word recognition. Examination of the correlations for both the forward and backward spans reveals that children who have longer WISC digit spans also display higher scores on all three word recognition tests. The correlations are all positive and reach statistical significance although the correlations of forward digit span with the three word recognition scores are somewhat larger than the correlations found for backward span.

Table I

	Simple Bivariate Correlations		Partial Correlations[a]	
	WISC Forwards Digit Span	WISC Backwards Digit Span	WISC Forwards Digit Span	WISC Backwards Digit Span
Closed Set Word Recognition (WIPI)	.42***	.28***	.25**	.12
Open Set Word Recognition (LNT-E)	.41***	.20**	.24**	.07
Open Set Word Recognition in Sentences (BKB)	.44***	.24**	.27***	.09

*** p < .001, ** p < .01

[a]Statistically Controlling for: Communication Mode Score, Age of Onset of Deafness, Duration of Deafness, Duration CI Use, Number of Active Electrodes, VIDSPAC Total Segments Correct (Speech Feature Perception Measure), Age

The right-hand portion of the table shows a summary of the partial correlations among these same measures after statistically controlling for differences due to: chronological age, communication mode, duration of deafness, duration of device use, age of onset of deafness, number of active electrodes and speech feature discrimination. As expected, when these seven "contributing variables" are statistically removed from the correlational analyses, the partial correlations between digit span and word recognition scores become smaller in magnitude overall. However, the correlations of the forward digit span with the three word recognition scores are still positive and statistically significant while the correlations with the backward digit spans are now much weaker and no longer reach significance. These results demonstrate that children who have longer forward WISC digit spans show higher word recognition scores and this relationship is observed for all three word recognition tests even after the other contributing sources of variance are removed.

In these results, forward digit span accounts uniquely for approximately 7% of the currently unexplained variance in the word recognition scores while backward digit span accounts for very little of the total variance in these scores. The present findings suggest a common source of variance that is shared between forward digit span and measures of spoken word recognition that is independent of other obvious mediating factors that have been found to influence variation in these outcome measures. As will be discussed further below, these findings are not overly surprising given that all three spoken word recognition tests require the use of some kind of working memory to maintain each lexical representation for a short period of time just prior to and during the child's response.

Digit Span and Sentence Duration

Although the findings on variation in digit span scores suggest that children who acquire language while using cochlear implant may differ both from normal-hearing children and amongst themselves in some basic information processing component, these data are not sufficient on their own to identify the basis for the observed differences. Additional converging measures of performance are needed to pinpoint the locus of these processing differences more precisely. Fortunately, another set of behavioral measures was obtained from these children for an entirely different purpose and these data were made available to us for several new analyses. These data consisted of a set of acoustic measurements of speech samples obtained from each child to assess speech intelligibility and to measure changes in articulation and phonological development following implantation (see Tobey et al. 2000). These speech samples provided a unique opportunity for us to use converging measures to further understand and explain the digit span results.

The speech samples consisted of utterances elicited using the stimulus materials and experimental procedures originally developed by McGarr (1983) to assess the speech intelligibility and articulation of deaf children. For the recordings made at CID, a clinician presented each child with meaningful English sentences using the child's preferred communication mode (either speech, or speech and sign), together with a printed version of the sentence on a large index card. All of the utterances produced by the children were originally recorded and stored digitally for playback to groups of naïve adult listeners who were asked to transcribe what they thought the children had said. From the duration measurements made by Dr. Tobey's research group of the twelve seven-syllable McGarr sentences, we were able to obtain measurements of the average time it took each child to produce each sentence.

These sentence durations provided us with quantitative measures of a child's articulation speed which is known from a large body of earlier research in the working memory literature to be closely related to speed of sub-vocal rehearsal (Cowan et al., 1998). Numerous studies with both adults and children over the past 25 years have demonstrated strong relations between speaking rate and working memory span for digits and words (see Baddeley, Thomson & Buchanan, 1975; Hulme & Tordoff, 1989; Johnston, Johnson, & Gray, 1987; Kail, 1992; Standing & Curtis, 1989). The results of these studies have been replicated with several different populations and suggest that measures of an individual's sentence duration reflect articulation speed; this measure, in turn, can be used as an index of rate of covert verbal rehearsal for phonological and lexical information in working memory (Baddeley, Thomson & Buchanan, 1975). Individuals who speak more quickly have been found to have longer memory spans than individuals who speak more slowly.

Several different explanations of the relationship between speaking rate and working memory span have been proposed in the literature. One account assumes that more forgetting occurs from immediate memory at slower speaking rates because fewer words can be articulated within a given interval of time (see discussion in Cowan & Kail, 1996). Another proposal assumes that the mechanism that controls speaking rate is the same one that regulates the speed of verbal rehearsal processes in short-term memory (Baddeley, 1992). Thus, more words can be maintained in working memory at faster rehearsal speeds. Regardless of which explanation is correct, the relation observed between measures of speaking rate and immediate memory span is a reliable and robust finding reported in the literature on working memory that has been found in several different populations of subjects.

The forward digit span scores for all of the 176 children tested at CID are shown in Figure 4 plotted against estimates of their speaking rates obtained from measurements of the seven syllable McGarr sentences. The digit spans are plotted on the ordinate; the average sentence durations are shown on the abscissa. The top panel shows mean sentence durations; the bottom panel shows the log-transformed mean sentence durations. Log-transformed scores were computed in order to obtain a more normally-distributed set of data. The pattern of results in both figures is very clear; children who produce sentences with longer durations speak more slowly and, in turn, have shorter forward digit spans. The simple bivariate correlations between forward digit span and both the raw and transformed measures of sentence duration were strongly negative and highly significant (r = -.55 and r = -.59; $p < .001$, respectively). For backwards digit span, the observed correlations were somewhat smaller, but still statistically significant (r = -.42 and r = -.42; $p < .001$).

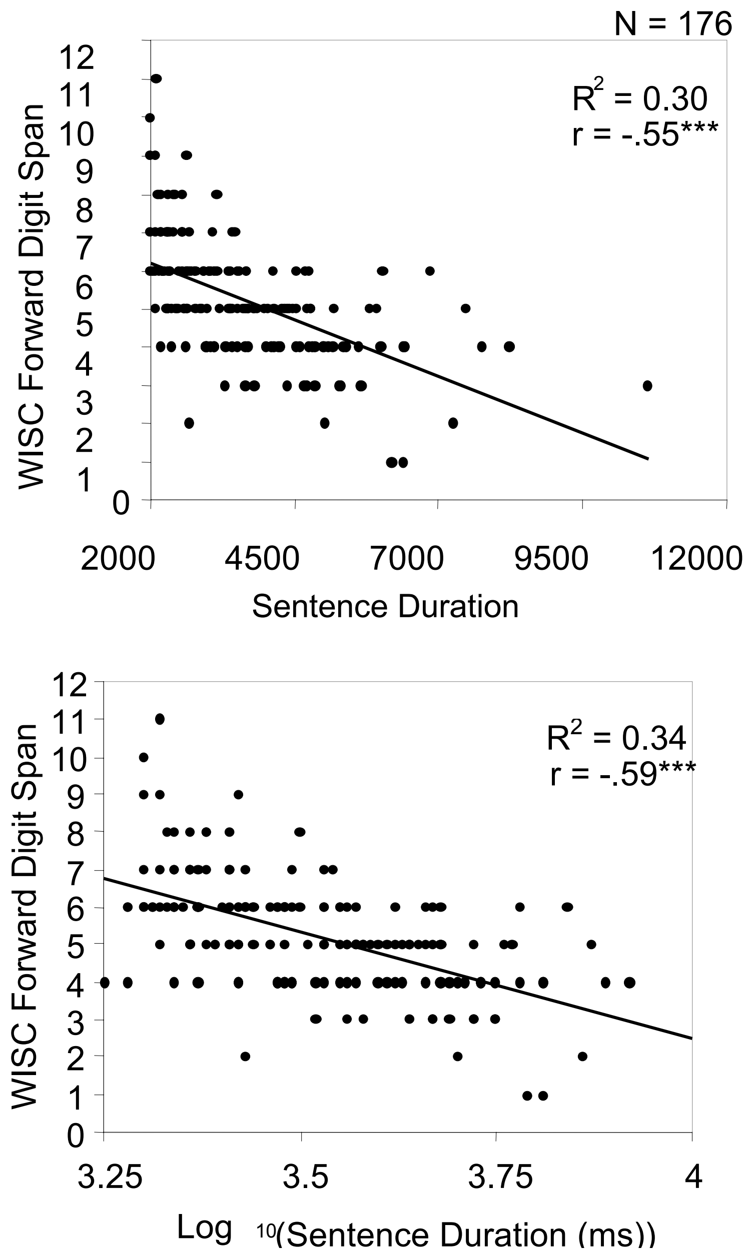


Figure 4. Scatterplots illustrating the relationship between average sentence duration for the seven-syllable McGarr Sentences (abscissa) and WISC forward digit span scored by points (ordinate). Each data-point represents an individual child. Non-transformed duration scores are shown in the top panel, log-transformed duration scores in the bottom panel. R-squared values indicate percent of variance accounted for by the linear relation.

These findings demonstrate that verbal digit span and articulation rate are correlated in this clinical population, as they are in normal-hearing school-age children and adults. That is, children who speak more quickly were found to have longer digit spans. This result suggests the existence of a common information processing mechanism responsible for the individual differences observed within both tasks, namely, limitations on verbal rehearsal speed.

Spoken Word Recognition and Sentence Duration

To determine if verbal rehearsal speed is also related to individual differences in word recognition performance, we next computed correlations between sentence duration and the three different measures of spoken word recognition described earlier. Table II shows the correlations between speaking rate and word recognition scores on the WIPI, LNT and BKB. Despite the fact that the sentence duration measure draws more heavily on speech production while the word recognition measures are designed to assess speech perception, the observed correlations are quite large. Table II also shows a summary of the partial correlations that were computed between the raw and log-transformed McGarr sentence durations and the three measures of spoken word recognition performance already described. As in the earlier analyses, differences due to possible mediating variables including traditional demographic factors, were once again statistically controlled for by using partial correlation techniques. In all cases, the negative correlations between sentence duration and word recognition remained remarkably strong and were highly significant.

The results of these correlational analyses demonstrate that slower speaking rates as measured by longer sentence durations are robustly associated with poorer scores on all three measures of word recognition, regardless of the response format of the test. Speaking rate accounted for approximately 25% of the currently unexplained residual variance in the word recognition scores even after the variability linked to other mediating variables was statistically controlled. These correlations are strong even for a word recognition test such as the WIPI which makes no apparent demands on overt speech production (recall that the child is only required to point to a correctly matched picture).

Table II

	Simple Bivariate Correlations		Partial Correlations[a]	
	Sentence Duration	Log (Sentence Duration)	Sentence Duration	Log (Sentence Duration)
Closed Set Word Recognition (WIPI)	-.65***	-.68***	-.47***	-.51***
Open Set Word Recognition (LNT-E)	-.60***	-.65***	-.45***	-.48***
Open Set Word Rec. in Sentences (BKB)	-.71***	-.76***	-.59***	-.64***

*** $p < .001$, ** $p < .01$

[a]Statistically Controlling for: Communication Mode Score, Age of Onset of Deafness, Duration of Deafness, Duration CI Use, Number of Active Electrodes, VIDSPAC Total Segments Correct (Speech Feature Perception Measure), Age

Why should processing speed play a role in what appears on the surface to be a relatively simple closed-set word identification task? Since spoken words extend temporally in time, early information must be retained as the remainder of the utterance is listened to and processed. In addition, some representation of the sensory pattern must be stored and maintained long enough for the listener to select the correctly matched picture or, in the open set tasks, to initiate a spoken/signed repetition of the item.

That is, even simple word recognition tasks such as the WIPI require some kind of memory representation to be maintained over a short period of time, and it is likely that covert verbal rehearsal is the processing mechanism used to maintain these representations as a response is arrived at and initiated.

Verbal rehearsal processes may be particularly important when spoken word recognition takes place under effortful or demanding conditions, such as those that exist for a child with a cochlear implant who is attempting to categorize a noisy and degraded auditory-only speech signal. If short-term memory is viewed as the "interface" to long-term memory, and noisy degraded signals induce listeners to attempt "top-down" contextual processing to recover the intended signal, the demands on rehearsal and maintenance may be considerable before a decision is finally reached. When identification is easy and the signal is well specified, identification is fast and individual differences in the ability to maintain a phonological representation may not figure as prominently (see Rabbitt (1968) and Pichora-Fuller, Schneider, & Daneman (1995), for related discussion).

Digit Span, Spoken Word Recognition, and Sentence Duration

The intercorrelations observed between digit span, articulation rate and spoken word recognition require further analyses in order to be fully interpretable. The high degree of intercorrelation among these three variables cannot be attributed to just a single source of variance (e.g., just working memory capacity or just verbal rehearsal speed) until we look at the correlations between each pair of variables with the other variable of interest statistically partialled out. More specifically, for each pair of variables (e.g., digit span and word recognition), it is necessary to determine if their correlation may be due entirely to their mutual relationship with the remaining variable of interest (e.g., sentence duration). The resulting partial correlations among the three variables are illustrated graphically in Figure 5. The correlations between digit span forwards, word recognition, and log-transformed sentence duration are shown in the top panel. The correlations between digit span backwards, word recognition, and log-transformed sentence duration are shown in the bottom panel.

As shown on the left side of each triangle, when sentence duration was partialled out of the analysis, the correlations between digit span and each of the three measures of word recognition essentially approached zero. This indicates that the associations observed between digit span and word recognition can be <u>entirely accounted for</u> in terms of individual differences in sentence duration, here interpreted as a measure of verbal rehearsal/processing speed. This is an interesting and important new result because it provides additional insight into the origins of the relationship between digit span and spoken word recognition. This finding also suggests that individual differences in verbal rehearsal speed may be largely responsible for the observed relationship between digit span and auditory word recognition, rather than individual differences in memory capacity.

Examination of the negative correlations shown at the base of each triangle indicate that the relationship between digit span and sentence duration remains fairly strong even when individual differences in spoken word recognition are statistically controlled for. These results are consistent with earlier studies suggesting that individual differences in verbal working memory as measured by an auditory digit span task can, in large part, be accounted for in terms of variation in speaking rate. Exactly why this is true is still a matter of current debate--perhaps limitations on speaking rate lead to forgetting during list output, or alternatively, perhaps a capacity limitation causes slowed and effortful production when capacity limits are stretched by the repeating-back of a many-syllabled sentence (see Cowan & Kail, 1996 for discussion). But the robustness of the relationship between digit span and sentence duration even when the spoken word recognition measures are partialled out also reassures us that the observed relationship is not likely a result of shared speech perception components in both tasks (perceiving the digits and sentences to be repeated). The relationship may be due to some other mediating factor.

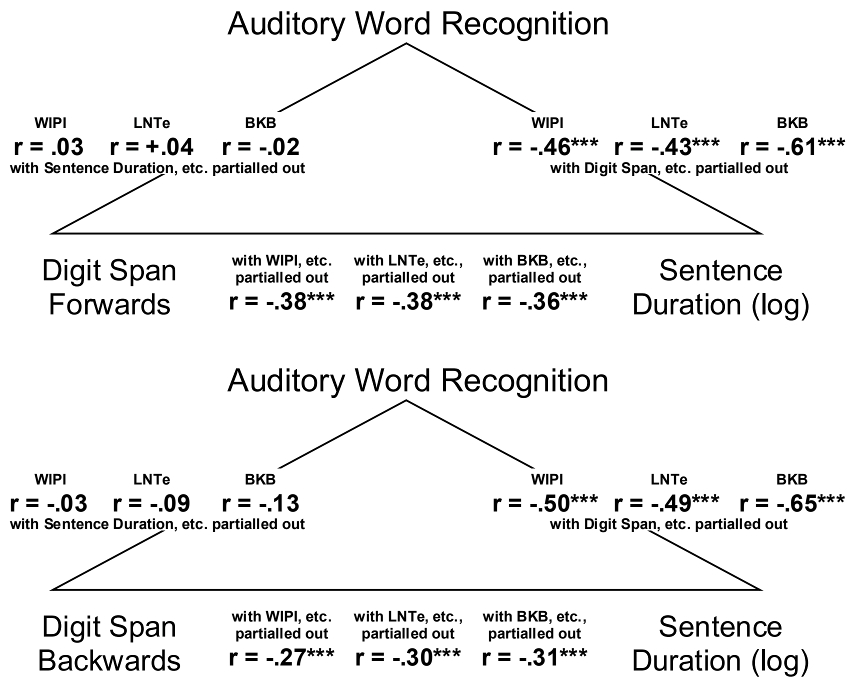


Figure 5. Illustrates the three-way relationship between auditory word recognition (WIPI, LNT-e, BKB), digit span, and sentence duration. The top panel shows the relations for forward digit span, the bottom panel shows the relations for backwards digit span. Communication mode, age at onset of deafness, duration of deafness, duration of implant use, number of active electrodes, chronological age at test, and VIDSPAC total segments correct (a speech feature perception measure), have all been partialled out of each correlation, along with the influence due to the variable listed at the opposite vertex.

Finally, the strong negative correlations shown on the right side of each triangle reveal that longer sentence durations are associated with poorer spoken word recognition performance even after individual differences in digit span are partialled out. From this asymmetric pattern of correlations in each triangle, we can conclude that a common feature of both digit span and sentence duration, probably best described as individual differences in immediate memory capacity, relates both tasks to word recognition, but that speaking rate incorporates an spoken auditory word recognition performance even after variability linked to differences in memory capacity is accounted for. Conceptualized in terms of a stepwise multiple regression analysis, this last statement is equivalent to saying that there is variability in the word recognition scores that is predicted only by the sentence duration measure and not by digit span scores.

Note also that essentially identical results were obtained for forwards versus backwards digit spans. The noticeable differences between forward versus backwards digit span in terms of their relationship to spoken word recognition performance are no longer evident once variability in verbal rehearsal rate is accounted for. The pattern of results that emerges from these analyses suggests that variation in performance on the examined measures can be traced to a common elementary process related to the speed of verbal rehearsal used to maintain phonological and lexical information in short-term working memory.

Discussion

Our investigation of working memory and speaking rate has provided new insights into the basic elementary information processing skills of deaf children with cochlear implants and the underlying cognitive factors that affect their speech and language performance on a range of outcome measures. These new studies were specifically designed to obtain process measures of performance that assessed the operation of verbal working memory in order to understand the nature of the capacity limitations in encoding and processing sensory information.

Several important findings have emerged from our analysis of the digit span and speaking rate data. The results obtained with these two process measures of performance suggest that working memory capacity and verbal rehearsal speed may contribute an additional unique source of variance to the outcome measures obtained with deaf children following cochlear implantation.

Although we found some overlap in the distributions of the digit span scores, the means of the forward and backward digit spans were shorter in length for the deaf children with cochlear implants than a comparison group of age-matched normal-hearing children. This pattern demonstrates clearly the presence of atypical development of short-term working memory capacity in these deaf children and supports our hypothesis that cognitive processing variables may contribute to explaining the variation and individual difference in a range of outcome measures used to assess speech and language performance in these children.

The presence of fundamental limitations in the capacity to process information in immediate memory— that is, to encode, maintain and retrieve verbal information in short-term working memory may have several important implications for other speech and language tasks as well. It is very likely that differences in information processing capacity and verbal rehearsal speed in immediate memory will propagate throughout the system and may cascade to higher levels of processing to influence performance on the behavioral tasks typically used to measure speech and language outcomes following implantation such as word recognition, vocabulary development, comprehension and even speech production.

The only demographic variable that was correlated with digit span and processing capacity was the child's communication mode. The deaf children who were immersed in oral-only environments displayed longer forward digit spans than the children who were in total communication environments. The presence of an effect of early sensory experience on forward digit span scores suggests that the stimulus environment and the specific kinds of interactions children have with their parents and caretakers in the language learning environment operate in a highly selective manner on a specific information processing mechanism and subcomponent of the human memory system that is used for initially encoding and maintaining phonological information in short-term memory. We suspect there may be something unique/different about the oral environment and the specific information processing activities that the child engages in on a regular basis that produces selective effects on the verbal rehearsal mechanism and the phonological coding of sounds.

Because children from TC environments may simply have less exposure to speech and spoken language in their early linguistic environment after implantation, they may display problems in both processing and rehearsing auditory information in short-term memory. In terms of initial encoding and recognition, the reduced exposure to speech and spoken language may affect the development of automatic attention and specifically the speed that speech signals can be rapidly identified and coded into phonological representations in short term memory. Thus, TC children may have problems in scanning and retrieving information from short-term memory. In terms of verbal rehearsal processes, TC children may have slower and less efficient verbal rehearsal processes once information finally gets into short-term memory simply because they have had less experience in producing speech and actively generating phonological patterns on output.

Passive exposure to speech without explicit analysis and conscious manipulation of phonological representations may not be sufficient to develop robust lexical representations of spoken words and fluency in control of speech production. Deaf children who receive cochlear implants may need to be actively engaged in processing spoken language in order to develop automaticity and automatic attention strategies that can be carried out rapidly without conscious effort or processing resources. This may be one important direct benefit of oral only education programs. The excellent spoken language skills acquired by children in oral-only programs may reflect the development of highly automatized phonological analysis skills which permit the child to engage in active processing strategies in perception that first involve decomposition of a speech pattern into a sequence of discrete phonological units and then the reassembly of those individual units into a sequence of gestures for use in speech production and articulation.

The development of phonological coding skills of this kind may result in increases in the speed and efficiency of constructing phonological and lexical representations of spoken words in short-term memory. Recovering the internal structure of an input pattern as a result of perceptual analysis and then reconstructing the same pattern in speech production may establish permanent links between speech perception and production and may lead to further development of highly efficient sensory-motor articulatory programs for verbal rehearsal and coding in working memory. Thus, the development of phonological processing skills may simply be a byproduct of the primary emphasis on speech and oral language skills in oral-only educational environments and may account for why oral children consistently display better scores on a wide range of outcome measures of speech and language, particularly oral language tests.

These new findings permit us to identify a specific information processing mechanism, the verbal rehearsal process in working memory, responsible for the limitations on processing capacity. Processing limitations are present in a range of behavioral tasks that make use of verbal rehearsal and phonological processing skills to encode, store, maintain and retrieve spoken words from working memory. We suggest that these fundamental information-processing operations are common components of almost all of the current outcome measures routinely used to assess both receptive and expressive language functions. The present findings suggest that the variability in performance on the traditional clinical outcome measures used to assess speech and language-processing skills in deaf children after cochlear implantation may actually reflect fundamental differences in the speed of information processing operations such as verbal rehearsal and the rate of encoding phonological and lexical information in working memory.

The present set of findings are theoretically significant because they provide new converging evidence from several different behavioral measures obtained on a large group of deaf children for the existence and operation of a common information processing mechanism used for storage and maintenance of phonological and lexical information in working memory. They also suggest a motivated explanation for the enormous variability and individual differences observed in a wide range of speech and language processing tasks that make use of a common set of verbal rehearsal processes. Verbal rehearsal is a fundamental processing component that is present in every one of the outcome measures typically used to assess speech perception, spoken word recognition, vocabulary, comprehension and speech intelligibility in this clinical population. As in normal-hearing children, differences in verbal rehearsal strategies may be the key to explaining the large individual differences in speech and language development observed in deaf children following cochlear implantation.

References

Adams, A.M., & Gathercole, S.E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders, 35*, 95-116.

Atkinson, R.C., & Shiffrin, R.M. (1968). The control of short-term memory. *Scientific American, 225,* 82-90.

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

Baddeley, A.D., Gathercole, S.E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation, Volume 8.* New York: Academic Press.

Baddeley, A.D., Thomson, N. & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575-589.

Balkany, T.J., Hodges, A., Miyamoto, R.T., Gibbin, K., & Odabasi, O. (2001). Cochlear implants in children. *Otolaryngologic Clinics of North America: Implantable Electronic Otologic Devices, State of the Art, 34,* 455-467.

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., Wright, M. Psarros, C., Rattigan, K., & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech Language & Hearing Research, 44*, 264-285.

Carpenter, P.A., Miyake, A., & Just, M.A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging, In *Handbook of psycholinguistics* (pp. 1075-1122), Gernsbacher, M.A. (Ed). San Diego, CA: Academic Press.

Cleary, M., Pisoni, D.B., Kirk, K.I., Geers, A., & Tobey, E. (2000). Working memory and language development in children with cochlear implants. Poster presented at *CI2000: The 6th International Cochlear Implant Conference.* Miami, Florida, February 2000.

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory & Language, 31,* 668-684.

Cowan, N., & Kail, R. (1996). Covert processes and their development in short-term memory. In S. E. Gathercole (Ed.), *Models of Short-term Memory* (pp. 29-50). UK: Psychology Press, Erlbaum Taylor & Francis.

Cowan, N., Wood, N.L., Wood, P.K., Keller, T.A., Nugent, L.D. & Keller, C.V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*, 141-160.

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

Engle, R.W., Kane, M.J. & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake and P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* Pp. 102-134. London: Cambridge Press.

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., Hitch, G.J., Service, E., Martin, A.J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*, 966-979.

Gazzaniga, M.S. (2000). *The New Cognitive Neurosciences, 2nd Edition.* Cambridge, Mass.: MIT Press.

Geers, A.E., & Brenner, C. (In press). Children implanted by age 5: Demographic and educational characteristics. *Ear & Hearing*.

Geers, A.E., Brenner, C., & Davidson, L. (In press). Factors associated with development of speech perception skills in children implanted by age 5. *Ear & Hearing*.

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).* St. Louis, MO: Central Institute for the Deaf.

Gupta, P. & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain & Language*, *59*, 267-333.

Haber, R.N. (1969). *Information-processing approaches to visual perception.* New York: Holt, Rinehart and Winston.

Hodges, A.V., Dolan-Ash, M., Balkany, T.J, Schloffman, J.J. & Butts, S.L. (1999). Speech perception results in children with cochlear implants: contributing factors. *Otolaryngology-Head & Neck Surgery, 12*, 31-34.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38,* 241-253.

Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47,* 72-87.

Johnston, R.S., Johnson, C., & Gray, C. (1987). The emergence of the word length effect in young children: The effects of overt and covert rehearsal. *British Journal of Developmental Psychology. 5,* 243-248.

Jusczyk, P.W., & Luce, P.A. (2002). Speech perception and spoken word recognition: Past and present. *Ear & Hearing, 23*, 2-40.

Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology, 28,* 899-904.

Kail, R., & Hall, L.K (2001). Distinguishing short-term memory from working memory. *Memory & Cognition. Special Issue, 29,* 1-9

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

Kirk, K.I., Diefendorf, A.O., Pisoni, D.B. & Robbins, A.M. (1997). Assessing speech perception in children. In L.L. Mendel & J.L. Danhauer (Eds.), *Speech Perception Assessment*. San Diego: Singular Press, Pp. 101-132.

Kirk, K.I., Pisoni, D.B., & Miyamoto, R.T. (2000). Lexical discrimination by children with cochlear implants: Effects of age at implantation and communication mode. In S.B. Waltzman & N.L. Cohen (Eds.), *Cochlear Implants* (pp. 252-254). New York: Thieme.

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.

Lachman, R., Lachman, J.L., & Butterfield, E.C. (1979). *Cognitive psychology and information processing: An introduction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Lindsay, P.H., & Norman, D.A. (1977). *Human information processing: An introduction to psychology, 2nd edition.* New York: Academic Press.

Locke, J.L. (1993). *The child's path to spoken language.* Cambridge, MA: Harvard University Press.

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.

Meyer, T.A., & Pisoni, D.B. (1999). Some computational analyses of the PBK Test: Effects of frequency and lexical density on spoken word recognition. *Ear & Hearing, 20,* 363-371.

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

Miyamoto, R.T., Svirsky, M.A. & Robbins, A.M. (1997). Enhancement of expressive language in prelingually deaf children with cochlear implants. *Acta Otolaryngologica, 117,* 154-157.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Pichora-Fuller, M.K., Schneider, B.A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America, 97,* 593-608.

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., Geers, A.E. & Tobey, E.A. (2000). Individual differences in effectiveness of cochlear implants in prelingually deaf children: Some new process measures of performance. *Volta Review, 101,* 111-164.

Pisoni, D.B. & Geers, A. (2000). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing. *Annals of Otology, Rhinology & Laryngology, 109*, 92-93.

Pisoni, D.B., Svirsky, M.A., Kirk, K.I., & Miyamoto, R.T. (1997). Looking at the "Stars": A first report on the intercorrelations among measures of speech perception, intelligibility, and language development in pediatric cochlear implant users. *Research on Spoken Language Processing Progress Report No. 21* (1996-1997). Bloomington, IN: Speech Research Laboratory, (pp. 51-91).

Posner, M.I. (1989). *Foundations of cognitive science.* M. I. Posner (ed.). Cambridge, MA: MIT Press.

Rabbit, P.M.A. (1968). Channel-capacity, intelligibility, and immediate memory. *Quarterly Journal of Experimental Psychology, 20,* 241-248.

Reitman, W.R. (1965). *Cognition and thought: An information-processing approach.* New York, Wiley.

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.

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.

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 Oto-Laryngologica, 117,* 755-759.

Standing, L. & Curtis, L. (1989). Subvocalization rate versus other predictors of the memory span. *Psychological Reports, 65,* 487-495.

Svirsky, M.A., Robbins, A.M., Kirk, K.I., Pisoni, D.B. & Miyamoto, R.T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science, 11,* 153-158.

Svirsky, M.A., Sloan, R.B., Caldwell, M. & Miyamoto, R.T. (2000). Speech intelligibility of prelingually deaf children with multichannel cochlear implants. *Annals of Otology, Rhinology, and Laryngology, 185*, 123-125.

Tobey, E.A., Geers. A.E., Morchower, B., Perrin, J., Skellett, R., Brenner, C., & Torretta, G. (2000). Factors associated with speech intelligibility in children with cochlear implants. *Annals of Otology, Rhinology and Laryngology Supplement, 185,* 28-30.

Wechsler. D. (1991). *Wechsler Intelligence Scale for Children, Third Edition (WISC-III).* San Antonio, TX: The Psychological Corporation.

Zwolan, T.A., Zimmerman-Phillips, S., Asbaugh, C.J., Hieber, S.J, Kileny, P.R. & Telian, S.A. (1997). Cochlear implantation of children with minimal open-set speech recognition skills. *Ear & Hearing*, 18, 240-251.

Speech Timing and Working Memory in Normal-Hearing Children and Deaf Children with Cochlear Implants[1]

Rose Burkholder

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

[1] This work was supported by NIH Research Grant DC00111 and Training Grant DC00012 to Indiana University. I extend my thanks Dr. Ann Geers and the Central Institute for the Deaf research staff for testing the cochlear implant users and Dr. Emily Tobey and the research staff at the Callier Advanced Hearing Research Center at the University of Texas, Dallas for data measurement. Also, special thanks are due for Miranda Cleary, Luis Hernandez, and Dr. David Pisoni for their help in establishing this study and Kara Kohnen for her help in data measurement.

Speech Timing and Working Memory in Normal-Hearing Children and Deaf Children with Cochlear Implants

Abstract. The relationship between articulation rate and working memory has been investigated extensively in normally developing children to understand how they access phonological information and rehearse it effectively through the articulatory control processes of working memory. Little, if any, research has examined articulation rate and working memory in children who receive highly degraded speech signals because of profound deafness. In the present study, 37 profoundly deaf children between 8 and 9 years of age who use cochlear implants (CI) were studied to measure their speaking rates, digit spans and assess the magnitude of the relationship between the two variables. Measurements from a group of normal-hearing (NH) age-matched children were also obtained to assess differences in speaking rate and digit spans between the two groups. Additional speech timing measures, including articulation rates, pause durations, and response latencies were obtained from recordings made during the recall portion of the WISC-III (Wechsler, 1991) auditory digit span task. The results showed that articulation rates, measured from sentence durations, were strongly related to immediate memory span in both NH and deaf children with CIs. In addition, pauses during recall were negatively correlated with memory span when both groups of children were considered together. These results replicate earlier findings on NH children reported by Cowan et al. (1994; 1998) and suggest that both subvocal verbal rehearsal speed, which is dependent on overt articulation, and serial scanning, which is carried out during recall pauses, contribute to immediate memory span in NH children as well as deaf children who have received CIs. The deaf children displayed significantly longer sentence durations and pauses during recall, and shorter digit spans compared to the NH children. This pattern indicates that slower subvocal rehearsal and scanning processes may contribute to their shorter digit spans. Such findings in a group of deaf children with CIs suggest that subvocal rehearsal and scanning processes are not just dependent on chronological development. Instead, in this clinical population, it appears that the absence of early auditory experience and phonological processing activities before implantation has measurable effects on the working memory processes that rely on verbal rehearsal and serial scanning of information in short-term memory.

Introduction

Current models of working memory generally include a component that is considered to act both as a storage medium and as a mechanism for rehearsal of phonological information (Baddeley, 1992). This component is called the *phonological* or *articulatory loop* and is a cognitive process that integrates speech-based information into working memory. The articulatory loop integrates phonological information into memory by using two components: the *phonological store* and the *articulatory control process*. The integration of information carried out by the phonological loop is assumed to occur through subvocal repetition or verbal rehearsal of phonological information. This repetition and verbal rehearsal occurs in a cyclic or looping fashion within the articulatory control process (Baddeley, 1986; 1992; Baddeley, Gathercole & Papagno, 1998).

Studies considering the repetition or rehearsal portion of the phonological loop have provided valuable information about the relationship between immediate memory span and vocal and subvocal repetition of items that are maintained in memory. It is generally accepted that items requiring less time for vocalization and, therefore, less time for subvocalization (Landauer, 1962) will be remembered more easily than longer words. Shorter words can be rehearsed faster than longer words. A faster rate of rehearsal is beneficial because it reduces memory decay. Memory decay ordinarily occurs if items are not refreshed within a two second time frame. This interval of two seconds has been suggested as the time limit of phonological working memory (Baddeley, 1986; Hulme & Tordoff, 1989; Standing, Bond, Smith & Isely, 1980).

The relationship between speaking rate and working memory in adult populations has often been examined by considering differences in word length and pronunciation rate. Word length is traditionally measured by the number of syllables in a word. Pronunciation rate is usually measured by how many syllables or words can be overtly articulated per second. Various spoken languages have been compared and assessed with regard to word length and pronunciation rate. Such comparisons have been useful in analyzing the differences in the memory spans observed in native speakers of different languages (Elliott, 1992; Naveh-Benjamin & Ayres, 1986; Powell & Hiatt, 1996).

A common method used to assess the relationship between word length, speaking rate, and the capacity of the phonological memory in individuals speaking different languages has been the digit span task. This task is used frequently because it is considered a reliable and valid way to tap into the mechanisms of working memory. Results of cross-language studies using the digit span task have shown that native speakers of languages that contain fewer syllables per digit name have longer memory spans (Elliott, 1992; Ellis & Hennelly, 1980; Naveh-Benjamin & Ayres, 1986). Again, shorter words are rehearsed more quickly. Within-subjects designs using bilinguals have shown that these speakers generally have longer digit spans in the language that has shorter digit names. These findings have been obtained even when experience and aptitude between the two languages have been taken into account (Chincotta & Underwood, 1996).

Other studies have also shown that shorter words are more easily recalled than longer words. This *word length effect,* similar to the bilingual digit span effect explained above, also occurs because shorter words require less time to articulate and subvocally rehearse than longer words (Baddeley, Thomson & Buchanan, 1975; Hulme & Tordoff, 1989). Research using nonwords has provided additional support for this proposal as well. These studies have shown that subvocal rehearsal of shorter words is faster, even in circumstances where context, meaning, and familiarity are stripped away from the words (Gathercole, Willis, Baddeley & Emslie, 1994; Hitch, Halliday & Littler, 1991). However, when articulatory suppression intended to prevent subvocal rehearsal has been included in these experiments, word length effects and language related digit span advantages are eliminated (Baddeley, Lewis & Vallar, 1984; Murray & Roberts, 1968). These findings provide support for the hypothesis that subvocal rehearsal plays a critical role in determining phonological memory span.

Other research has focused on the effects of presentation rate on memory recall. It is generally accepted that faster rates of presentation produce better recall of phonological information (Mackworth, 1964; Murray & Roberts, 1968). This research also examines the effects of self-presentation rate, or overt articulation speed, on working memory. However, differences in articulation rate are not likely to be dependent on the characteristics of the language but rather on attributes of the speakers. Individuals that speak more rapidly than others, when given the same speech stimuli, generally display longer memory spans in verbal short-term memory tasks (Cowan et al., 1998; Elliot, 1992). These findings suggest that individual variability may also play a role in controlling overt speaking rates and immediate memory span.

In addition to individual differences, developmental effects have been associated with variability in speech rate (Cowan, 1992; 1994; Cowan et al., 1998; Hulme & Tordoff, 1989; Kail & Park, 1994). As children develop, their articulation rates increase. This increase in speaking rate likely contributes to an increase in covert verbal rehearsal speed. Enhanced covert rehearsal speeds will then, in turn, contribute to an increase in immediate memory span.

The positive relationship between articulation rate and working memory span is a reliable finding in the literature. Memory span can be predicted linearly from measures of overt speaking rates of words (Baddeley et al., 1975; Baddeley, 1992) and nonwords (Hulme, Maughan & Brown, 1991) in adults and children (Hulme & Tordoff, 1989). Several recent studies of children have shown that these memory span predictions are based on the maximal rate at which children can repeat various lists of words (Cowan et al., 1994; Hulme & Tordoff, 1989; Kail, 1997). Repetition tasks using word lists are commonly used to analyze children's articulation rates and to examine these rates' influence on children's memory spans. The results of these tasks have generally demonstrated that there is a positive correlation between speaking rate and memory span.

However, developmental differences in the relationship between speaking rate and memory span have been found when children of different age ranges are compared. One example of these striking differences was a negative correlation found between speeded articulation rates and working memory in four-year-old children (Cowan et al., 1994). In this study, Cowan and his colleagues used rapid repetition of word pairs to assess the speaking rates of four- and eight-year-old children. As was expected, eight-year-old children displayed the typical relationship between speaking rate and memory span. That is, eight-year-old children able to speak faster had longer memory spans. In contrast, the opposite relationship was observed in the four-year-old children studied. This finding is surprising because children at this age are assumed to be in the early stages of developing subvocal rehearsal strategies (Flavell, Beach & Chinsky, 1966; McGilly & Siegler, 1989). Such counter-intuitive results suggest that the influence of speaking rate on working memory may depend on development. These results also suggest that the relationship between speaking rate and immediate memory should be studied more extensively to establish more consistent conclusions about its role in the development of speech and language skills.

The relationship between speaking rate and working memory capacity in children is of particular importance because of the implications it has for language acquisition and other areas of cognitive development, such as reading. Strong links have been found between phonological processing, reading, and language in NH children (Kail & Hall, 1994; Williams, 1984; Zifcak, 1981), as well as in deaf children (Daneman, Nemeth, Stainton & Huelsmann, 1995). At all ages and hearing abilities, working memory is likely to be quite important for language processing. First, working memory serves to maintain sensory information, making immediate language comprehension possible. Second, working memory stores and organizes additional information, allowing language production to occur effectively (Cowan, 1996).

In several studies, Gathercole and her colleagues (1994) reported a direct causal relationship between phonological working memory skills and vocabulary acquisition in children age 4 and 5. At this time in development, the knowledge and the ability to utilize subvocal rehearsal strategies begins to emerge (Flavell, Beach & Chinsky, 1966; McGilly & Siegler, 1989). Such co-occurring phenomena accentuate the importance of the relationship between verbal rehearsal rates, working memory, and milestones in cognitive development.

In addition to traditional measures of speaking rate, other speech timing measures obtained through different methodologies have been examined for possible influences on working memory. Research examining speaking rates and working memory in children has been expanded to include additional speech timing measures obtained during the actual recall portion of memory span tasks (Cowan, 1992; Cowan et al., 1994; 1998). These new measures include individual articulation times and interword pause durations, as well as the preparatory intervals preceding list recall. Like pre-test or non-recall based measures of speaking rate, speech timing measures taken from the actual recall process have contributed new insights into the relationship between temporal characteristics of speech and working memory.

In a study examining speech timing measures during recall, Cowan et al. (1994) suggested that interword pause times may be crucial determinants of phonological working memory. The importance of pause times in recall is based on the assumption that they reflect the operation of retrieval or scanning processes during recall (Cowan, 1992). This mechanism is taken to be developmentally linked because older children have been found to have shorter pause durations than younger children (Cowan et al., 1998). Additionally, Cowan et al. (1994; 1998) and other researchers examining memory development have shown that older children have longer memory spans than younger children. Cowan et al. (1998) also found that children with shorter interword pauses had longer memory spans than their peers.

Taken together, the recent findings by Cowan et al. (1994; 1998) suggest that the memory span increase in older children might be associated with shorter interword pauses during serial recall tasks. These shorter interword pauses, according to Cowan, demonstrate that scanning mechanisms are being executed faster and more efficiently in the older children. Additionally, from the same body of research, Cowan and his colleagues observed that interword pauses increased as children reached their span-limiting list length. This is the point at which the digit span task is assumed to be the most cognitively demanding. This result suggests that pauses during recall might play an important role in the memory retrieval process during development by serving as an index of the serial scanning period.

The shorter interword pauses observed in older children with larger memory spans may reflect the maturation of covert rehearsal and scanning mechanisms. This maturation may result in more effective and rapid serial scanning processes in older children than in younger children (Cowan, 1999; Hale, 1990). This factor, along with increases in articulation speed, may enhance the ability to engage in efficient memory recall as children develop. These findings have led Cowan and colleagues (1998) to conclude that there may be two processing methods used by children that contribute to working memory capacity. These two processes are serial scanning or retrieval and subvocal rehearsal of phonological information. Not only are these processes important in memory development per se, they are also important because they appear to become operative during different stages in development. Such asynchronies during development illustrate the intricate organization and acquisition of working memory and attention abilities in children (Cowan et al., 1998; Cowan, 1999).

Prior to Cowan’s work examining speech-timing measures during serial recall, most studies that have investigated speech rates and working memory in children have used the same basic methodologies. The similarities in previous methodologies have caused some concern that there may be additional dimensions of this relationship yet to be found (e.g., pause time, as suggested by Cowan, 1999). In the majority of research that preceded studies measuring speech timing during recall, children’s speech timing was evaluated in speeded articulation tasks. Speeded articulation tasks are designed to obtain maximal articulation rates by requiring participants to vocalize lists of words, letters, or numbers as fast as they can (Cowan et al., 1994; Hitch, Halliday & Littler, 1989; Hulme & Tordoff, 1989; Kail, 1997). Although there has been concern expressed that repeating lists may be confounding because of memory demands imposed by the task (Henry, 1994), it is still a common method of measuring the articulation rates of children (Ferguson, Bowey & Tilley, 2002).

Few research designs outside of a recall task have measured articulation rates of children by using more natural yet still experimentally controlled methods. Such methods previously used with adults, like reading short passages (Naveh-Benjamin & Ayres, 1986) or meaningful sentences, may be applicable to developmental populations as well. These methodologies may be useful because they provide a more consistent and representative measure of speaking rate exhibited in children on a regular basis. Maximal speaking rates, in contrast, may never be revealed in the same manner outside the laboratory setting. In addition to being ecologically valid, non-speeded articulation designs would be useful to determine the generality of research that has linked speeded articulation rates to working memory span.

To our knowledge, there have also been few research designs examining the relationship between speaking rate and measures of working memory in clinical populations of children. Developmentally delayed children with mental handicaps are one population in which links between phonological processing and working memory have been examined. Early research on this population suggested that atypical verbal rehearsal and encoding strategies contribute to digit span recall disadvantages (Ellis & Anders, 1969). However, other more recent research indicates deficiencies in central executive functioning (Conners, Carr & Willis, 1998). Unfortunately, such conclusions concerning executive or verbal rehearsal deficits in this population are likely confounded by other factors related to cognition and intelligence.

In order to avoid confounds related to cognition and intelligence, developmental populations that exhibit normal intelligence yet have an articulatory or phonological disadvantage for other reasons should be studied. Children with specific language impairment (SLI) meet these criteria. Significant amounts of research have shown that children with SLI exhibit deficits in working memory (Gathercole & Baddeley, 1990; Leonard, 1998; Sussman, 1993). However, such deficits are not thought to be related to problems that characterize SLI (Cowan, 1996). Similarly, working memory deficits in SLI children are not thought to be associated with sensory processes or speech sound discrimination (Gathercole & Baddeley, 1990; Sussman, 1993). However, it would be interesting to examine rehearsal and working memory in a clinical child population in which overt and covert rehearsal capabilities may be attenuated due to a deficit in speech discrimination and articulation. Profoundly deaf pediatric cochlear implant users manifest these characteristics ideally making them a particularly suitable population in which to study covert rehearsal and working memory.

The speech of the deaf has been studied for a number of years because of its relevance to the quality of communicative abilities in deaf individuals. One distinguishing characteristic of deaf speech is its reduced rate of articulation. Reduced speaking rates have been found in deaf individuals prior to the availability of cochlear implants (Nickerson, 1975), as well as in cochlear implant candidates (Leder et al., 1987). In addition to speaking rate, intelligibility of deaf speech is another common assessment used to evaluate the quality of language production in deaf individuals (McGarr, 1981; 1983; Osberger, Maso & Sam, 1993; Osberger, Robbins, Todd & Riley, 1994).

The intelligibility of deaf speech refers to how well short speech samples can be understood by naïve, normal-hearing, adult listeners. The McGarr Sentence Intelligibility Test (McGarr, 1981) was one of the first instruments constructed to assess the speech intelligibility of deaf speakers, and was specifically designed to evaluate the intelligibility of the speech in deaf children. McGarr (1983) found that deaf children display atypical speech production strategies and often fail to provide the same speech information that is ordinarily provided by normal-hearing adults.

A more recent study using the McGarr Sentence Intelligibility Test has found that the quality of speech production exhibited by deaf children using cochlear implants is closely linked to their speech intelligibility (Tobey & Hasenstab, 1991). Additionally, it has been shown that decreased speech intelligibility in cochlear implant users may be related to their slow articulation rate. Using the McGarr Sentence Intelligibility Test, Pisoni and Geers (1998) found that the intelligibility of speech was related to the speed at which it was articulated by children using CIs. This relationship indicates that the longer the duration of the McGarr sentences, the less intelligible they were to naïve normal-hearing listeners who were asked to transcribe them.

These results suggest that there are communicative advantages for pediatric cochlear implant users who are able to articulate faster. One such advantage is simply being more intelligible than their slower speaking peers. An additional advantage fast-talking CI users is that they may be more capable of planning and maintaining their speech output in memory with less effort. Such decreased working memory demands during speech planning likely improve the fluidity and clarity with which speech output is articulated.

One factor that is important to both articulation rate and intelligibility of speech in children using cochlear implants is the nature of the early sensory and linguistic experience used to develop these abilities. Within the CI population, communication environments can vary measurably according to what kind of strategies are utilized by each child. Communication strategies used by deaf children with CIs vary across a continuum ranging from exclusive oral communication (OC) to total communication (TC), a method utilizing oral communication supplemented with manual sign and lip reading strategies. By assessing where children fall on this continuum, a classification into either the OC or TC group is usually made. This classification method has allowed for comparisons of the two groups on a variety of communicative and cognitive measures based on the nature of the early auditory and linguistic experiences of the children.

In their studies of speech intelligibility of OC and TC children, Osberger et al. (1993; 1994) found that the most intelligible deaf children used oral communication. Based on the relationship of intelligibility to speaking rate, it follows that OC users speak faster than their TC counterparts. An ability to speak faster and more intelligibly in OC children may be a direct result of their early communicative experiences after implantation. The most beneficial early experiences that the OC users have are oral-aural activities. In addition to encouraging the ability to produce speech, oral-aural activities provide the necessary auditory feedback to deaf children using CIs. Auditory feedback is crucial for these children because it provides a mechanism for them to self-monitor and improves their speech output and intelligibility.

In addition to displaying higher speech intelligibility than TC users, OC users have also been found to have longer phonological working memory spans exhibited by significantly longer WISC-III (Wechsler, 1991) forward digit spans. This finding suggests that their phonological working memory capacities and/or capabilities may be directly benefiting from increased articulation rates (Pisoni, 1999). More direct support for this relationship has been shown in a recent study in which CI users with faster articulation rates had longer WISC-III digit spans (Pisoni, Cleary, Geers & Tobey, 2000). Given this relationship between articulation rate and working memory in children, the working memory advantage displayed by OC children may be related to both overt articulation and covert verbal rehearsal abilities; both of these abilities are crucial to a phonological memory task such as digit span recall.

The CI children can be expected to utilize covert rehearsal strategies because verbal rehearsal strategies have been measured in deaf children without cochlear implants (Bebko, 1984; Liben & Drury, 1977). In addition, it has been shown that deaf children, like their NH peers, display word length effects in memory tasks (Campbell & Wright, 1990). More importantly, in a study examining verbal and spatial working memory in a sample of deaf children using CIs, Cleary, Pisoni and Geers (2001) found evidence of verbal rehearsal and encoding in the CI users. In some cases, the verbal rehearsal strategies of the children with CIs were as efficient as in a control group of NH children. Based on these findings, it is reasonable to expect that CI users in the present study will utilize covert rehearsal as well. Since covert rehearsal can be assumed, it follows that those CI users able to utilize covert rehearsal at a faster pace because of faster overt articulation rates will have longer memory spans. In addition, we expect that the fastest speaking NH children will have the longest memory spans.

The present study was designed to expand on the results described above and to complete the missing link currently observed in the relationship between speech timing measures and working memory in CI children. To establish such links, both speaking rate and verbal memory retrieval timing were examined in relation to working memory in a CI sample. Such examinations were conducted on the CI group as a whole and separately according to communication mode. Articulation rate measures were made by examining sentence durations from a non-speeded McGarr sentence repetition task. Speech timing measures were made by examining articulation and pause times embedded within WISC-III digit span recall. In addition to seeking replication of the relationship between articulation rates and WISC-III digit spans in pediatric CI users, this study further examined this relationship in a group of age-matched NH children. The magnitude of the relationship between articulation rate and working memory in each group of children was compared to determine how the characteristics of the relationship vary between the two different populations.

The comparison between speaking rate and working memory performance in pediatric cochlear implant users and NH controls seems appropriate based on earlier research comparing the memory capabilities of deaf and NH children. This earlier research has shown, not surprisingly, that there are large differences in phonological memory performance between deaf children and their NH age-matched peers. In one study examining phonological memory in deaf and NH children, Banks, Gray and Fyfe (1990) found that deaf children had more difficulties recalling details read in written text. In phonological memory tasks dependent on sequential information similar to the digit span task, deaf children have also been found to lag behind NH children (Waters & Doehring, 1990).

These results have been corroborated by studies of juvenile CI users. In a recent study, Cleary, Pisoni and Geers (2001) found that deaf children using CIs had significantly shorter working memory spans for both verbal and spatial patterns than do NH children. In addition, it has been determined that CI users have shorter forward and backward digit spans than their NH peers (Pisoni, 1999; Pisoni et al., 2000). The present investigation also seeks to replicate the finding that CI children have shorter working memory spans than their NH peers. Because of the close relationship between working memory spans and articulation rate, we also expect that deaf CI users will have longer McGarr sentence durations than a group of NH age-matched peers. In addition, we predict that within the CI group, TC users will have longer sentence durations than the OC users. Longer sentence durations are also expected to co-occur with shorter digit spans in the TC group.

Finally, for the first time in a CI population, speech-timing measures were obtained from the recall portion of the WISC-III auditory digit span task. To assess the timing of the spoken recall, response latencies, durations of individual articulations, and interword pauses in WISC-III digit span lists were obtained from both the CI and NH groups. These measures were evaluated according to what influences they exert on recall capabilities, paying close attention to the role of pause durations. Based on the recent findings reported by Cowan (1999), using a similar methodology with NH children, both articulation rate, obtained from the McGarr sentence task and the interword pauses obtained from the spoken recall of the WISC digit span task should be related to memory spans in the current NH group of children.

Of greater interest in this study is how these two processing rates function in the CI group. The importance of these two timing measures stems from the general hypothesis that articulation rate reflects the rate of subvocal verbal rehearsal and that pause durations in recall reflect the time spent scanning and retrieving information from short-term memory. Close evaluation of the CI children's speech timing measures were made to determine if their relationship to digit spans are equivalent to that observed previously in NH children and the current NH control group. The relationship between speech timing and memory span is considered in order to determine how it influences the digit spans of CI and NH children and within the TC and OC users. We predict that both of the processes reflected in speech timing, subvocal rehearsal and serial scanning are atypical in the CI population, particularly the TC users. These differences are expected to be observable through decreased articulation rates in McGarr sentences and longer interword pauses during the recall portion of the WISC-III forward digit span task.

Method

Participants

Thirty-seven deaf 8- to 9-year-old children (*M* = 8.70, *SD* =.51) who use cochlear implants were studied. Twenty-five of the children were male and 12 were female. The deaf children were tested at Central Institute for the Deaf (CID) in St. Louis, Missouri as part of a larger ongoing study called "Cochlear Implants and Education of the Deaf Child" (see Geers, 2000). Most of the subjects had congenital profound hearing loss. Five of the children lost their hearing after birth, between the ages of 9 and 18 months (*M* = 14.00, *SD* = 4.58). The average age of onset of deafness for all children was approximately two months of age. Implantation of the device occurred between 1.72 and 5.03 years (*M* = 3.04, *SD* = .88). The duration of deafness before implantation ranged from 0.60 to 5.03 years (*M* = 2.88, *SD* = 1.13). The duration of implant use for this group of children ranged from 4.46 to 6.87 years (*M* = 5.66, *SD* = .64). Prior to their inclusion in the CID study, the deaf children were evaluated through intelligence testing to ensure that they fell within reasonable limits expected for their appropriate age range. Only children that met these criteria were tested at CID and included in the present study.

The CI users were classified into two different communication groups based on whether they used primarily oral (OC) or total communication (TC). Total communication refers to a training mode utilizing manual sign and lip reading strategies, in addition to speech. The classification into TC or OC groups was based on scores assigned to the children just prior to implantation and three consecutive years after implantation. Additionally, their communication training programs were evaluated at the time of testing. The scores used in this evaluation ranged from "1," signifying a program primarily stressing the use of sign and lip reading (generally in the form of signed exact English or cued speech, not American Sign Language (ASL)) to "6," representing an oral-only regime. Each score assigned at each year of evaluation was then summed producing communication mode scores that could range from 5 to 30. This summed score determined the mode of communication that the CI children had most consistently used for a four-year period and at testing. Children with summed scores of 15 and below were considered to be TC users. Those children with scores above 15 were considered to be OC users. This method of division is based on the original scoring scale in which the lower scores (1-3) most accurately represent total communication and the higher scores (4-6) most accurately represent oral communication.

The actual range of scores obtained by these children was between 6 and 30 (*M* = 18.92, *SD* = 7.32). Children classified into the OC mode were determined to have been communicating primarily through oral communication during the four years prior to testing and at the time of testing. Children determined to be communicating orally with the supplement of manual sign and lip reading during the four years prior to testing and at testing were considered to be using TC strategies. Twenty-two children fell into the classification of OC users while the remaining 15 were considered to be TC users. All CI children were administered the Wechsler Intelligence Scale for Children (WISC-III) (Wechsler, 1991) forward and backward digit span tasks, the McGarr Sentence Intelligibility Test, and a variety of speech perception and comprehension tests.

A comparison group of 36 age- and gender-matched NH children was also included in this study (*M* = 8.75, *SD* = .69). An independent sample test of the mean ages of the control and CI group showed no difference in the ages of the children, ($t(71) = -.399$, $p = .691$). The NH group of children consisted of 24 males and 12 females. All children were reported by their parents to be monolingual native speakers of American English. Parental report also indicated that the children had no known speech, hearing, or attentional disorders at the time of testing.

This group of children was determined to be normal hearing by results of a brief hearing screening given by the researcher prior to beginning the experimental procedure. Using a standard portable pure-tone audiometer (Maico Hearing Instruments, MA27) and TDH-39P headphones, each child was tested at tone pulses of 250, 500, 1000, 2000, and 4000 Hz at 20 dB in first the right ear and then the left ear. All testing of the normal-hearing children was done in a small, quiet testing room at the Speech Research Laboratory that was equipped with a closed-circuit television camera. The NH children were also administered the WISC-III digit span and McGarr Sentence Intelligibility tests. Additionally, the NH children, but not the CI children, were administered a speeded articulation task and the Peabody Picture Vocabulary Test- Third Edition, Form A (PPVT-III; Dunn & Dunn, 1997).

Stimuli and Materials

The McGarr Sentence Intelligibility Test that was completed by both the NH and CI using children included a set of visual stimuli used to elicit short sentences from the children. A set of 36 McGarr sentences (McGarr, 1981) were printed in 36 point Times New Roman font and each sentence was affixed to a three by five inch note card. The 36 sentences included 12 each at 3-, 5-, and 7- syllables. The utterances of the sentences spoken by both groups of children and the NH children's utterances from all other tasks were recorded onto digital audiotape (Sony Walkman TCD-D8) via a uni-directional headset condenser microphone (Audio-Technica ATM75). This apparatus did not physically or mechanically interfere with the deaf children's usage or placement of their cochlear implant.

Auditory stimuli used in a speeded articulation task administered only to the NH children were recorded digitally by the experimenter, a female native speaker of American English, from a similar geographical region as the participants. Audio recordings were made using the Speech Acquisition Program (SAP; Hernandez, 1995) in a sound booth. All stimuli were equated for amplitude as a batch to 70 dB. The stimuli were transcribed by four NH listeners to insure that they were highly identifiable in isolation before they were used. The stimuli for the rapid articulation task were four pairs of digits (1-6; 2-9; 7-4; 5-8) and four pairs of words (book-glove; car-spoon; fish-pig; leaf-egg) (Kail, 1997). Each pair was recorded as a separate token and presented to the children individually over a high quality tabletop loudspeaker (Advent AV570).

Additional testing materials were used to obtain vocabulary measures from all children and speech perception measures from the CI users. The PPVT (Dunn & Dunn, 1997) was given to the NH children. The Test of Auditory Comprehension of Language-Revised (TACL-R; Carrow-Woolfolk, 1985) was administered to the CI children. The CI children were also tested using the open-set spoken word identification Lexical Neighborhood Test (LNT) for easy (LNTe), hard (LNTh), and multisyllabic words (mLNT) (Kirk, Pisoni & Osberger, 1995). The Word Intelligibility by Picture Identification (WIPI; Ross & Lerman, 1979) provided a means for testing closed-set spoken word identification in the CI users. Sentence perception was measured in the CI group by administering the open-set Bamford-Kowal-Bench Sentence List Test (BKB; Bench, Kowal, & Bamford, 1979). Speech-feature discrimination was evaluated using the VIDSPAC. VIDSPAC is a video game test of speech contrast perception, specifically designed for use in hearing-impaired children (Boothroyd, 1997). All performance tests for the deaf children were also administered at Central Institute for the Deaf as part of the larger, ongoing study.

Procedure

Digit Span Task. The WISC forward and backward digit span test was administered to both the deaf and hearing children. The CI children were administered the task using live voice presentation, with lip reading available, from a trained clinician at CID. Following standard administration procedures, one digit per second was read from the list by the experimenter. There were two lists at each length. List lengths of the forward digit span task began with two digits and increased to a maximum of nine digits. List lengths of the backward digit span task began with two digits and increased to a maximum of eight digits. Two practice lists were also administered in the backward digit span task. Testing concluded when both lists at the same length were incorrectly recalled or not attempted by the child. The task was administered in the same way to the NH children by the experimenter in the Speech Research Laboratory at Indiana University in Bloomington, Indiana. The entire administration procedure of the digit span task was recorded in both groups of children.

Analog audiotape recordings of the deaf children's digit span responses were made via a lavaliere clip-on microphone worn by the clinician during administration. The sessions were originally recorded in order to verify that the digit presentation rate was approximately one digit per second. The presentation rate was verified to be consistent through examinations made by a research assistant at the Speech Research Laboratory.

The analog recordings of each deaf child's WISC-III digit span response were digitized and stored separately as ".wav" sound files in the CoolEdit Pro Limited Edition (LE) (Syntrillium Software Corporation, 1996) digital waveform editing program. These responses were used in this study to obtain the speech timing measures of articulation rates, response latencies, and pause durations within the spoken digit span responses. During the digitizing process, the recordings were sampled at 44.1 KHz with a 16-bit resolution. Forty-five CI children were originally recorded and digitized in this manner. However, eight children were later eliminated from the study. The recordings were judged to be too poor to measure accurately from a visual waveform. The digit span responses of the NH children were also all digitized and segmented into separate lists and stored using CoolEdit Pro LE in the same manner as the CI children's recordings. Once recordings were digitized, they were measured using CoolEdit LE to determine the latencies of response and articulation and pause durations in the verbal recall portion of the task.

The acoustic measurements made on all of the children's usable recordings of each list of digits included responses latencies, articulation times, and pause times. All measures were made in seconds to the nearest millisecond using simultaneous waveform and spectrogram views. Measurement was conducted in CoolEdit Pro LE by selecting beginning and end points of the desired speech or pause segment with a computer mouse cursor. Response latencies were measured from the end of the clinician's or experimenter's concluding utterance in a list to the initiation of the first digit uttered in the child's response. Any response preceded by extraneous utterances by a child was not included in the analysis of response latency. If a child began to verbally recall the list before the experimenter was done administering it, response latency measures were also disregarded. However, articulation time and pause duration measures were still made on these responses.

Individual articulation times were measured for each digit uttered in each list by locating the start and finish of the vocalization of the digit. Pauses were measured similarly from the end point of a digit to the beginning of the next digit in the list. The individual measures made within one list were averaged to give the mean individual interword pauses and mean individual articulations in lists of 2, 3, and 4 digits. Articulation and pause measures within each list were also summed to give a total articulation time and pause duration time. In addition, all articulations and pauses were included in one measure of the entire utterance duration. The average of each measure was calculated if two lists at one length were correctly recalled and measured. (See Figure 1 for a schematic representation of the measuring points that were made on the digit span lists.)

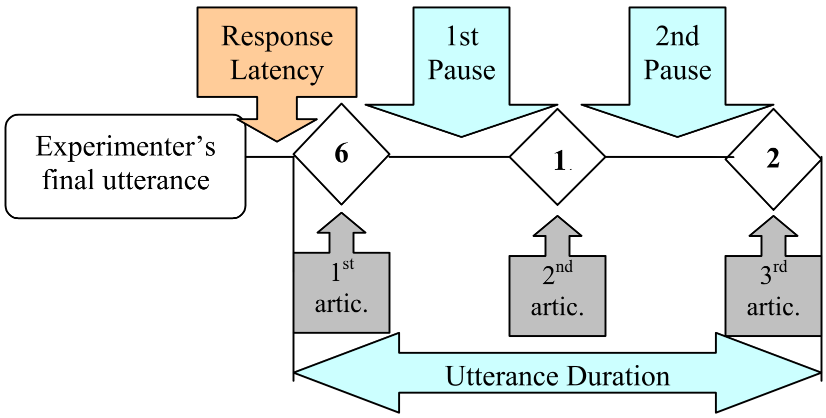


Figure 1. Schematic representation of speech timing measures made on WISC-III digit span responses. Example of list length three (6 1 2).

Only the measurement data from correctly recalled lists were used in the final analysis. Any measurements made on incorrect lists or lists with additional vocalizations or repetitions of correct numbers were disregarded. Correct responses from two practice items preceding the backward digit span task were also measured. Although all responses meeting these criteria were measured for the CI group, measurements of the NH children's digit span recall were only made up to lists of digit length four in both the forward and backward task. This limitation was made because few CI children could progress past lists of length four. Therefore, making most measures in lists longer than length four was unnecessary for the NH children. However, additional measures were made and considered at the list length limit (the longest list correctly recalled) for both groups of children. Recordings were measured by the primary researcher and a research assistant to determine inter-rater reliability. Correlations between the two rater's measures were determined to be between .88 and .97 when all the measures of response latencies, articulation durations, and pause durations were considered separately. The primary researcher's measurements were used in the final analysis.

McGarr Sentence Repetition Task. Both NH and CI children were presented with the 36 McGarr sentences in verbal and printed forms and asked to repeat them in their "best speaking voice." Sentences were presented randomly by shuffling the index cards with the sentences' written text prior to testing. The clinician or experimenter first read a sentence and then placed the index card with a printed version of the sentence in front of the children. The clinician also manually signed the sentences to the CI users if they required it. Access to lip reading was also available to all children.

Upon seeing the sentence to be spoken, the children were asked to reproduce the sentence in their best speaking voice. For the CI children, the quality of the speech recorded was closely monitored during testing. If the clinician noted any incomplete or incorrect portions of the sentences, the child was asked to repeat it up to a maximum of three times. This procedure was followed in order to elicit the best speech sample possible from the CI children.

Digital audiotape recordings were made of the utterances from both groups of children completing the McGarr Sentence Intelligibility Test. The sentences spoken by the NH children were digitized and stored as separate tokens in CoolEdit Pro LE. Duration measurements of the entire spoken sentences were then made on each group. The average durations of sentences at each syllable length (3, 5, 7) and the average total duration of all sentences were calculated for the two groups. The measurements of the CI group were completed at Callier Advanced Hearing Research Center at the University of Texas, Dallas, in cooperation with CID. The measurements of the NH group were completed at the Speech Research Laboratory using CoolEdit Pro LE.

In addition, the sentence durations of the 36 NH participants used in this study were compared to the durations of another group of 26 age- and sex-matched NH children whose data were collected at CID. This comparison was made to address the issue of possible testing effects caused by different speakers administering the test to the CI and NH children. Comparisons of the two groups of NH children showed no differences in speaking rate at syllable lengths 3 and 5 and at all lengths averaged overall. However, at syllable length 7, the children tested in the Speech Research Laboratory were found to speak at a significantly faster rate ($p = <.05$). See Figure 2 for all NH average McGarr durations. As a whole, these results show that the speaking rates between the two groups of NH children are fairly consistent. This finding was desirable because it provides some confidence that the speaking rates of children completing the McGarr repetitions were not globally influenced by the test administrators' speaking rates.

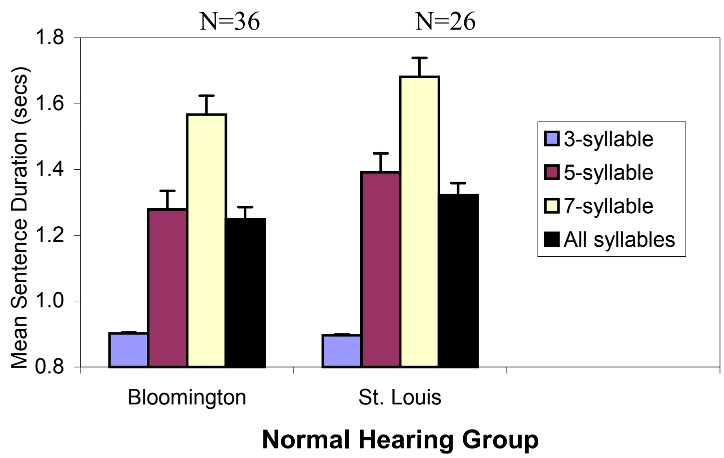


Figure 2. McGarr sentence duration comparison in two normal-hearing groups of children. Error bars represent the standard error.

After examining the distributions of the durations of the CI children and the NH children tested in Bloomington, one NH and one deaf child were eliminated from the final data analyses involving speaking rate. These children, both males, were excluded because their average McGarr sentence durations deviated significantly from the mean at all syllable lengths. For example, at syllable length seven, the NH child, the fastest speaker in the group, was more than two standard deviations below the mean when the average of the 7-syllable sentences measured in seconds was calculated ($M = 1.06$, $z < -2$). The CI user eliminated was the slowest speaking ($M = 8.23$, $z > 3$) and was an OC user. The decision to eliminate the CI user was made independent of the communication group classification.

Rapid Articulation Task. Normal-hearing participants were told that they would hear pairs of two words or two numbers together and that they needed to repeat them as quickly as possible, in pairs, until they were told to stop. The pairs were presented to all the children in the same order through the tabletop speaker at 70 dB SPL. Number pairs were presented first followed by the word pairs. Each child was prompted by the experimenter to stop after completing ten repetitions of each of the pairs. The children were permitted to rest between each set of pairs repeated if desired.

The audio recordings of the rapid articulation of word and number pairs were then digitized and segmented into their own sound files and stored in CoolEdit Pro LE where they were later measured. Measurements were made to determine both the duration of the entire repetition sequence and the rate of articulation. The average duration of each utterance, measured in seconds, was calculated for each child. From this measurement, the number of words uttered per second was determined, serving as another measure of articulation rate.

Results

WISC Digit Span Scores. Previous results regarding digit spans in CI and NH children were replicated in the present study. As we expected, CI children displayed shorter WISC-III digit spans than their NH peers. Additionally, TC users showed shorter forward digit spans than OC users. These results suggest that the CI children, particularly those using TC programs, are atypical in their phonological working memory abilities as indexed by traditional digit span measures.

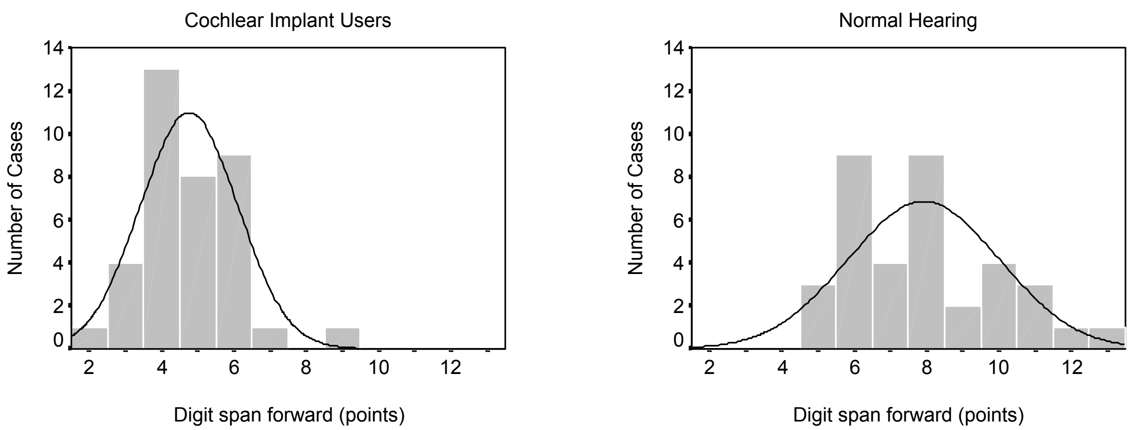


Figure 3. Distribution of forward digit span scores in CI and NH groups.

Digit span scores reflect the number of lists correctly recalled, not including practice items of the backward digit span condition. A point was awarded for each list correctly repeated to obtain the digit span scores. The range of possible scores on the forward digit span task is 0 to 16. The possible scores in the backward task range from 0 to 14. The significant difference ($t(59.43) = -7.71$, $p < .001$) in forward digit span between the NH ($M = 7.9$, $SD = 2.09$) participants and the CI users ($M = 4.8$, $SD = 1.34$) was just over 3 points. The distributions of these digit span scores are shown in Figure 3.

In addition, NH children ($M = 4.6$, $SD = 1.25$) had longer backward digit spans ($t(64.30) = -3.86$, $p < .001$) than the CI users ($M = 3.2$, $SD = 1.80$). Figure 4 illustrates the distribution of the scores for each group on backward digit span. Within the CI group, there were only significant differences between the forward digit spans ($t(35) = 2.19$, $p = .035$) of the 22 OC ($M = 5.1$, $SD = 1.32$) and 15 TC ($M = 4.2$, $SD = 1.21$) users. The distribution of forward digit span scores in the CI groups is shown in Figure 5. Backward digit spans between the TC ($M = 3.1$, $SD = 1.73$) and OC ($M = 3.3$, $SD = 1.88$) groups were nearly identical ($t(35) = .229$, $p = .821$) as shown in Figure 6. Figure 7 illustrates the mean score of all groups in the forward and backward digit span condition.

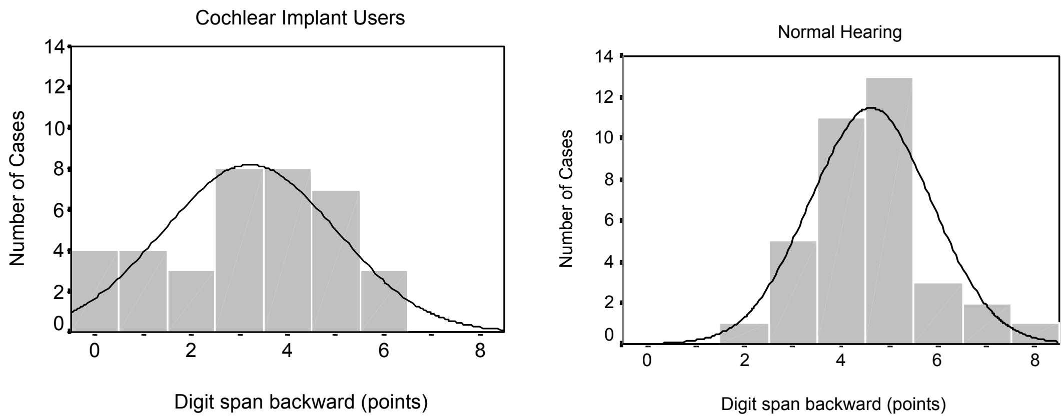


Figure 4. Distribution of backward digit span scores in CI and NH groups.

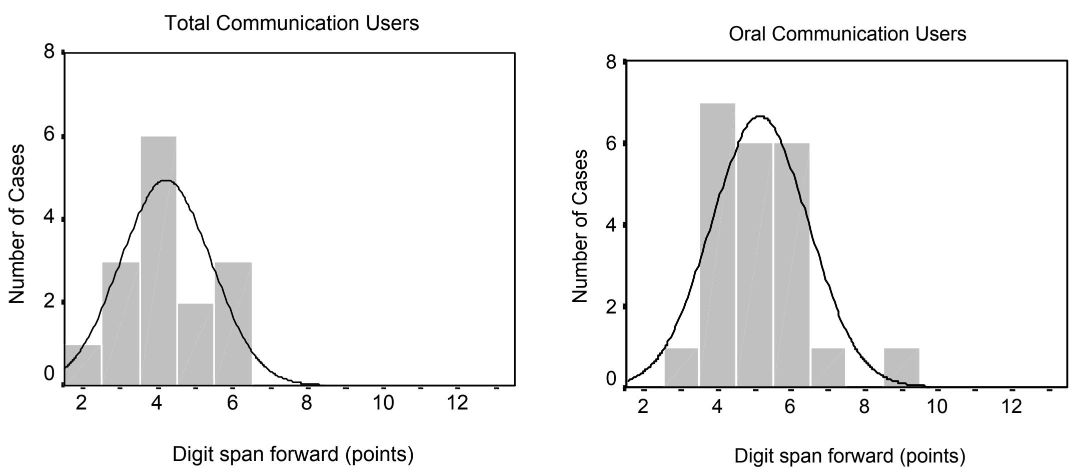


Figure 5. Distribution of forward digit span scores in TC and OC groups.

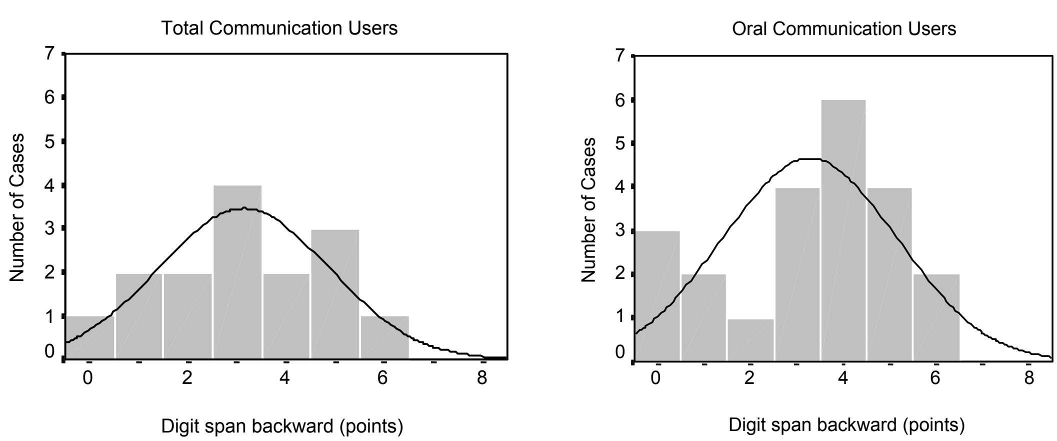


Figure 6. Distribution of backward digit span scores in TC and OC groups.

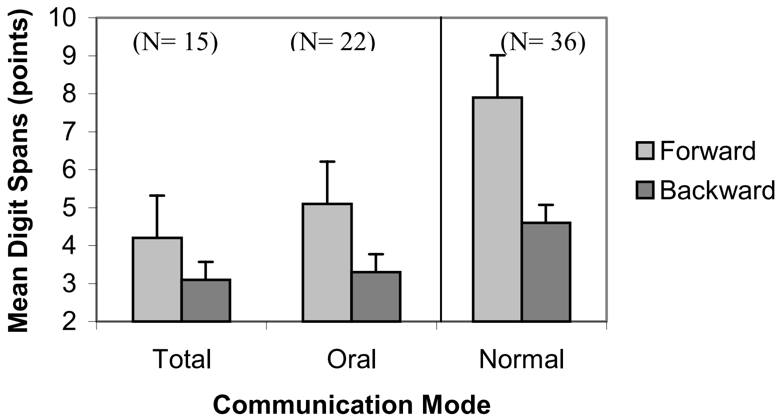


Figure 7. Average forward and backward WISC-III digit span scores in points. Error bars represent the standard error.

Limiting Span Measures. In addition to the conventional scoring system of the digit span task, all participants were evaluated based on their maximum span or memory limiting list length (Cowan et al., 1994; Cowan, 1999). The limiting span length is the longest list that could be recalled correctly in the task. At the maximum list length, it is assumed that children are at their information processing capacity where the task is most cognitively taxing. Obtaining this measure for every child provides an opportunity for a comparison of performance when each child is most challenged with the task and at the capacity of their immediate memory span.

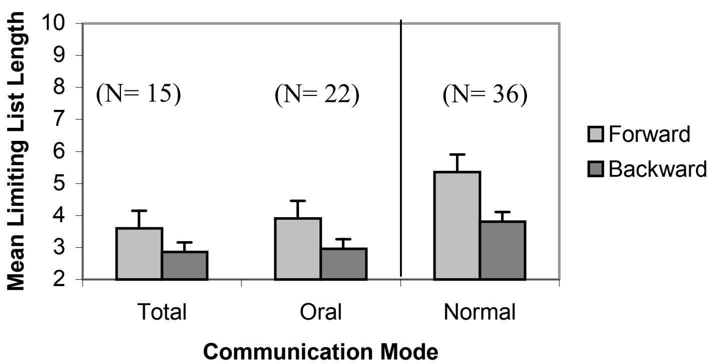


Figure 8. Average forward and backward WISC-III digit span limiting list lengths. Error bars represent the standard error.

Consistent with the point-based scoring system, we also observed differences in limiting list length between the NH and CI children. NH children had longer span limiting list lengths on average in both the forward ($M = 5.36$, $SD = 1.22$) and backward ($M = 3.81$, $SD = .749$) conditions ($t(57.81) = -6.62$, $p < .001$) than the CI children did in the forward ($M = 3.78$, $SD = .750$) and backward ($M = 2.92$, $SD = 1.18$) conditions ($t(70) = -3.82$, $p < .001$). However, there were no significant differences between the maximum list length recalled by the OC and TC groups in either the forward ($t(35) = 1.239$, $p = .223$) or backward ($t(35) = .238$, $p = .813$) condition. In fact, the mean limiting list length of the forward condition was nearly the same in the OC (*M* = 3.91, *SD* = .750) and TC (*M* = 3.60, *SD* = .737) groups, although the OC users had a slight advantage. The limiting list length of the backward digit span tasks were also slightly longer in the OC (*M* = 2.96, *SD* = 1.29) group than in the TC (*M* = 2.86, *SD* = 1.03) group. Figure 8 shows the means of the limiting list lengths of each group relative to the others.

McGarr Sentence Durations. As predicted, we observed significant differences in the speaking rates of all the groups at each of the three syllable lengths of the McGarr sentences. A post-hoc analysis utilizing Tukey's HSD procedure showed that NH children speak the fastest at all syllable lengths, TC children the slowest, and OC children display intermediate levels of speaking rate. The durations of all groups were significantly different from each other at all syllables lengths and overall when all syllable lengths were considered together, according to post-hoc analysis. Figure 9 illustrates these differences.

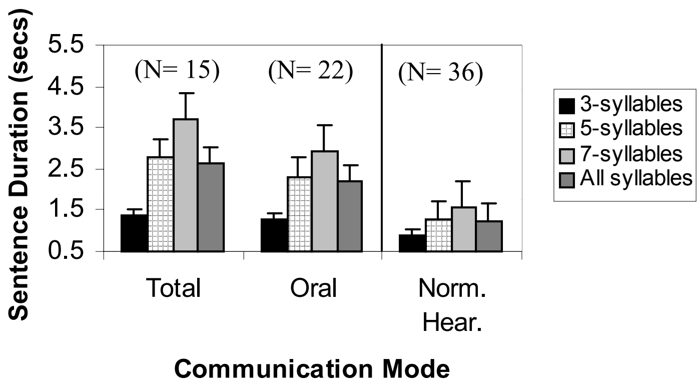


Figure 9. Mean McGarr sentence durations. Error bars represent the standard error.

Consistent with previous studies examining the relationship between working memory and speaking rate, the McGarr sentence durations were found to be negatively correlated with WISC-III forward digit spans in both the CI and NH groups. Syllable length seven of the McGarr sentences was chosen as the best measure of speed of articulation because the greater amount of syllabic content it provided allowed for more variance within the groups. In both the CI and NH groups, spoken durations of the McGarr sentences at syllable length seven were related to WISC-III forward digit spans using Pearson product correlational analysis. For this analysis, a natural log transformation of the raw durations that were measured in seconds was used. This transformation was necessary to normalize the slightly skewed raw data.

Hearing Ability	Log McGarr 7-Syllable Duration
Deaf DS Forward	-.516**
DS Backward	-.629**
NH DS Forward	-.369*
DS Backward	-.036

Communication Mode	Log McGarr 7-Syllable Duration
Oral DS Forward	-.376
DS Backward	-.648**
Total DS Forward	-.693**
DS Backward	-.712**

Tables 1-2. Relationship of McGarr 7-syllable sentences and WISC-III forward and backward digit spans in the CI and NH groups and the OC and TC groups. *$p < .05$, **$p < .01$

In the entire CI group, speaking rate was negatively correlated with forward digit spans. However, in the OC group, the correlation between speaking rate and forward digit spans failed to reach significance (p = .077). In addition, the correlation between backward span and speaking rate was strong in both CI groups but nonexistent in the NH group. Only the CI children showed a relationship between articulation rate and backward digit span while the relationship was absent in the NH children. See Tables 1 and 2 for the values of these correlations.

Partial correlations between the average McGarr 7-syllable duration and WISC-III digit spans were conducted on the CI group to control for the possible influences that speech perception, word recognition and language abilities may have on speaking rate. Three separate partial correlations were carried out controlling for scores obtained on a closed-set word identification task (WIPI), an open-set task of sentence repetition (BKB) and a test of speech feature discrimination (VIDSPAC). To control for language comprehension related to intelligence, the scores of an auditory language comprehension test (TACL-R) were also partialled out of the correlation. Table 3 shows a summary of the correlations that resulted from partialling out these measures individually.

Partialled out Variable	**Digit Span Condition**	**Log McGarr 7 Durations**
VIDSPAC	Forward	-.490**
Speech feature discrimination	Backward	-.526**
WIPI	Forward	-.362*
Word identification	Backward	-.443**
TACL Age	Forward	-.285*
Auditory language comprehension	Backward	-.450**
BKB	Forward	-.403*
Open set sentence repetition	Backward	-.399*

Table 3. Correlations of McGarr 7-syllable sentence durations and WISC-III forward digit spans in CI children with speech perception and comprehension measures partialled out of analysis.
*$p < .05$, **$p < .01$

The strengths of the correlations between speaking rate and digit span were reduced somewhat after these analyses. However, the overall relationship between speaking rate and digit spans in the CI group still remained strong and significant. Chronological age was not related to either digit span or speaking rate in any of the groups. Therefore, no adjustment was made to control for this factor in either the CI or NH groups.

Speeded Articulation Task. Although the speaking rates obtained using the McGarr Sentence Intelligibility Task were related to forward digit span in the NH group ($r = -.37$, $p < .05$), there was no relationship between the speeded articulation of either the word pairs ($r = -.21$, $p = .251$) or the digit pairs ($r = -.25$, $p = .164$) and digit span. Figure 10 shows scatterplots of these relationships. Despite the finding that speeded articulation and digit span did not correlate, the speeded articulation rates of word pairs did correlate with the McGarr sentence durations at all syllable lengths. This correlation is not surprising given that both repetition tasks involved similar speech stimuli. However, the speeded articulation of digit pairs was only related to the McGarr sentence durations at syllable length seven. This indicates that the children's speaking rate performances, relative to the rest of the group, were fairly consistent in both a speeded and non-speeded articulation task. These correlations are shown in Table 4. Figure 11 shows the relationship of the speeded articulation rate of word and digit pairs to the mean McGarr durations at syllable length seven, the measure taken to represent non-speeded articulation in this study.

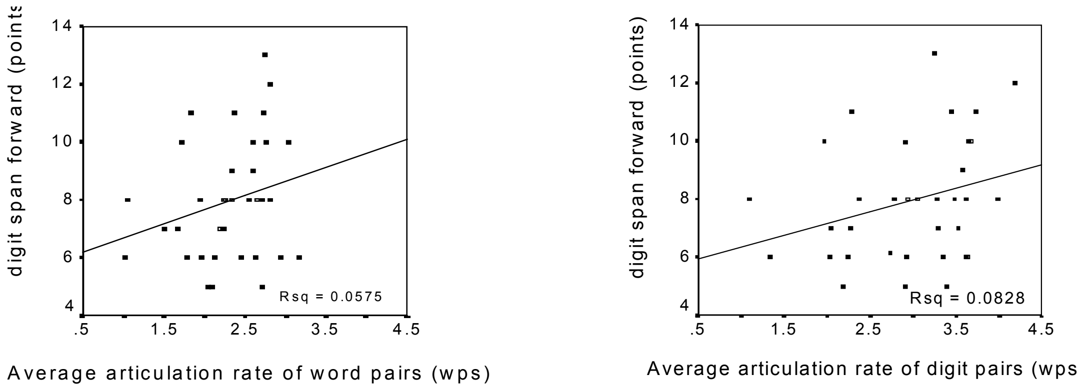


Figure 10. Correlation between rapid articulation rate and WISC-III forward digit span scored in points in the NH children.

Syllables *Stimulus Pairs*	3	5	7	All McGarr
Words	-.338*	-.359**	-.468**	-.410*
Digits	-.212	-.312	-.349*	-.311
Both	-.250	-.327	-.399*	-.346*

Table 4. Correlations between mean rapid articulation rate of words and digits with McGarr sentence durations for the NH children. *p < .05, **p < .01

Speech Timing Measures During Digit Recall: Articulation Durations. For the analysis of the speech timing measures during recall, only the responses from the digit span forward condition were considered and reported here. Analysis of the speech timing measures obtained during recall revealed no differences between the three groups in the average articulation of individual digits in any of the list lengths (2 (F (2, 66) = .262, p = .771), 3 (F (2, 68) = .689, p = .506), and 4 (F (2, 55) = 1.005, p = .373)) or the limiting list lengths (F (2, 68) = .818, p = .446) considered from the digit span forward condition. No relationship was found between the average articulations and digit span forward at any of the list lengths or the limiting list length when all children were considered together or when evaluated in groups according to hearing ability or communication mode.

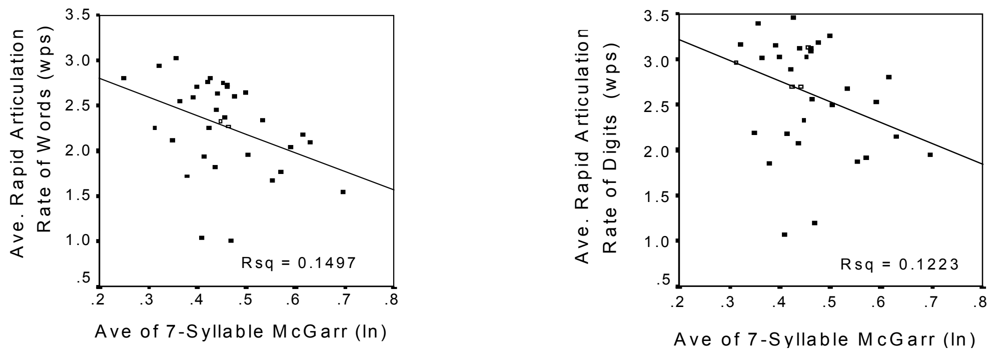


Figure 11. Correlations between mean speeded articulation rate of word and digit pairs and McGarr 7-syllable sentences in NH children.

Speech Timing Measures During Digit Recall: Response Latencies. The average response latencies of all the correct forward digit span lists did show a weak negative relationship (r = -.294, p = .014) with forward digit span, scored in points, when both the deaf and NH children were considered together. This relationship is plotted in Figure 12.

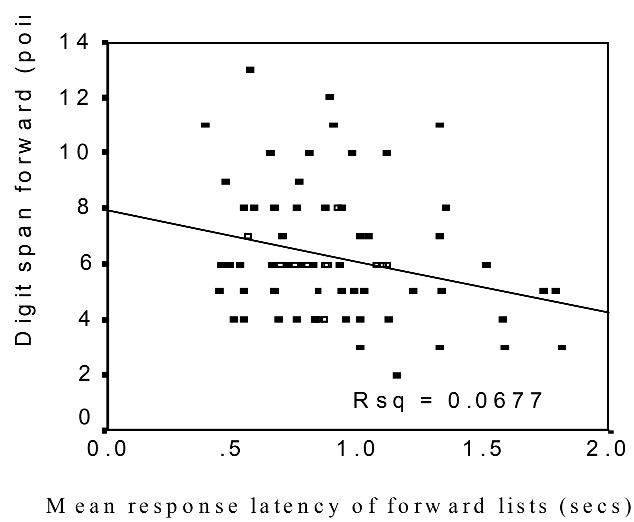


Figure 12. The correlation between mean response latency of all correctly recalled WISC-III forward digit span lists and digit span forward scored in points in both deaf and NH children.

However, this relationship appears to be driven by the TC group only. The correlation between the average response latencies of the forward lists and digit span forward points in the TC group was significant (r = -.599, p = .023). There was also a strong negative correlation between the average forward response latencies and the limiting list length in the TC group (r = -.722, p = .004). Scatterplots illustrating this relationship are shown in Figures 13 and 14.

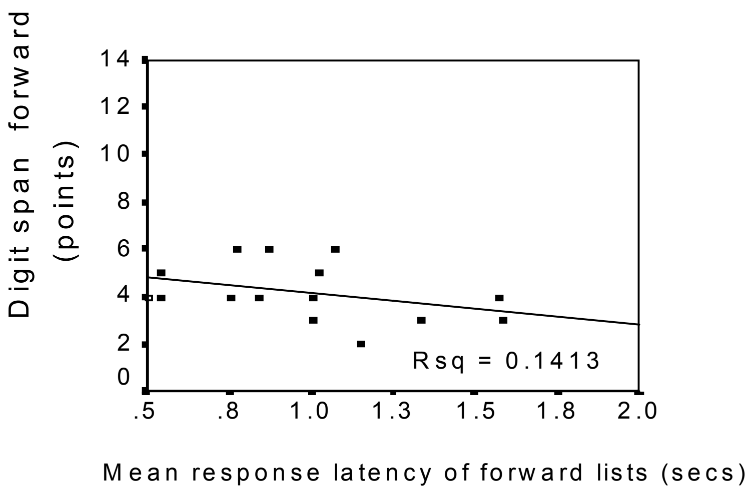


Figure 13. Correlation between mean response latency in digit span forward lists and digit span in TC cochlear implant users.

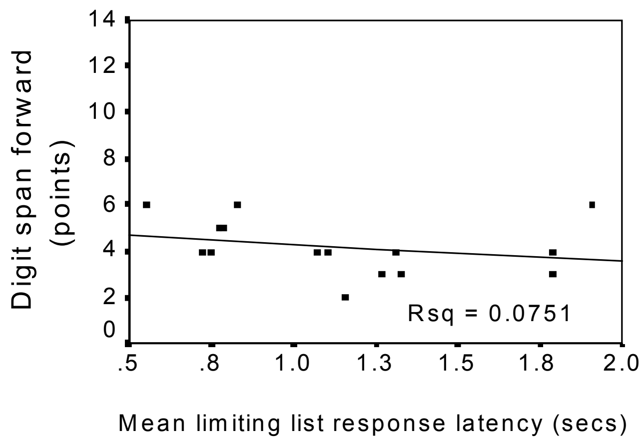


Figure 14. Relationship of mean response latencies in limiting list lengths to WISC-III forward digit span in TC cochlear implant users.

A significant correlation between the average response latency at the list limit and the length of the list limit was found in the NH group (r = .346, p = .045), although it is the inverse of the relationship observed in the TC group. See Figure 15 for an illustration of this relationship. Although the relationship of response latencies and forward digit span was particularly salient in the TC children and an unexpected relationship between span limiting response latencies and span length was found in the NH group, there was no difference in the mean response latencies observed between the three groups in the forward lists or at the span limiting list.

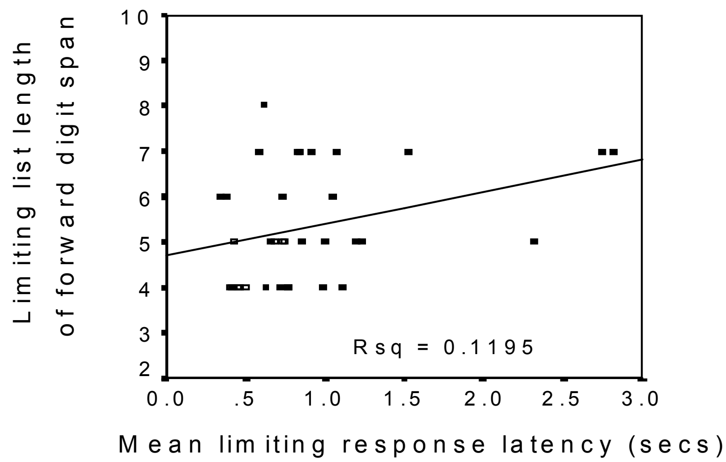


Figure 15. Relationship of mean response latency at forward limiting list length and the limiting list length of WISC-III digit span forward in normal hearing children.

Speech Timing Measures During Recall: Pause Durations. One speech timing measure that was shown to display differences among the groups was average pause duration. The average of individual pauses taken during recall in the forward condition were longer in both of the CI groups than in the NH children at list lengths three ($F(2, 66) = 18.583$, $p < .001$) and four ($F(2, 59) = 15.261$, $p < .001$). In addition, the average pauses taken from each child's own limiting list length were longer ($F(2, 68) = 17.109$, $p = .000$) in the TC ($M = .518$, $SD = .312$) and OC children ($M = .459$, $SD = .241$) than in the NH children ($M = .175$, $SD = .157$). Within the CI group, post-hoc analyses showed there was no difference between OC and TC users in the average pause durations at any forward list length, although there was a tendency for the pauses taken by the TC users to be longer than those taken by the OC users. See Figure 16 for an illustration of the average pause lengths in all groups at lists with three and four digits and at the limiting list length.

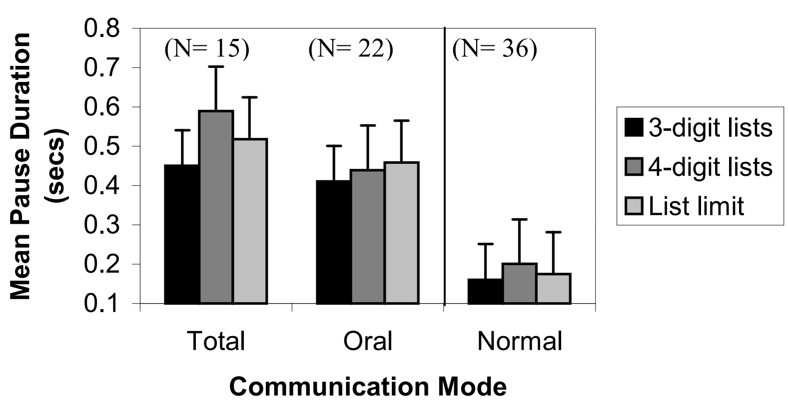


Figure 16. Average single pause durations during WISC-III forward digit span recall. Error bars represent the standard error.

The correlation of pause durations in forward digit lists of length three and four with forward digit span was found to be significant when the entire sample of children was considered together. See Table 5 for the values of these correlations. Figure 17 is an example illustration of the relationship between mean pauses and forward digit span in the group of all 72 children. This figure plots the mean interword pause taken during the limiting list length against digit span forward. There was no correlation between pause duration and digit span when each group was considered separately.

List Length	Digit Span Forward
2 Digits	-.234*
3 Digits	-.412**
4 Digits	-.323*
Span List	-.403*

Table 5. Correlations between mean pause duration during WISC-III forward digit span recall and forward digit span scores in all children. *p < .05, **p < .01

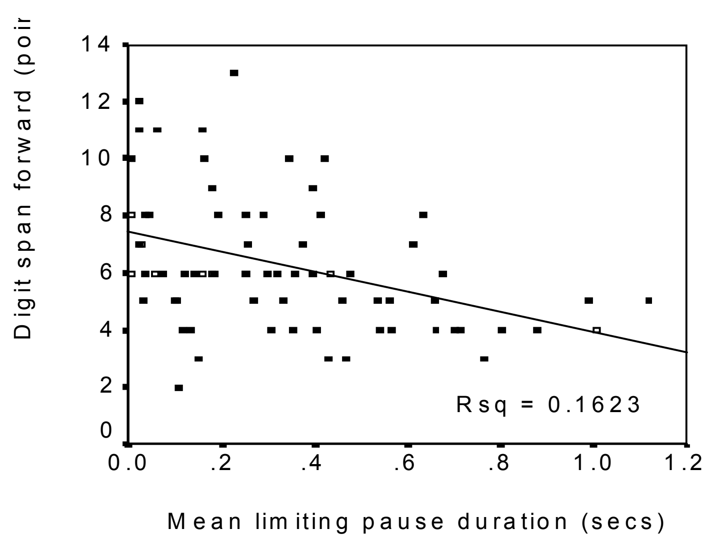


Figure 17. Correlations between mean interword pause duration in span limiting list and WISC-III digit span forward in all children.

Discussion

As expected, the results of this study replicated previous studies showing that profoundly deaf children with cochlear implants have shorter digit spans and slower speaking rates than their NH peers. Additionally, within the CI group, TC users displayed slower speaking rates and shorter forward digit spans than did OC users. These results provide additional support for the proposal that speaking rate and working memory are closely related in this clinical population. It is assumed that slower speaking rates reduce or impede subvocal verbal rehearsal processes and, therefore, the maintenance of phonological information in working memory. Both of these differences are probable causes for the shorter digit spans observed in the CI group.

More direct confirmation of this hypothesis was obtained correlationally. However, bi-directional influences between the measures of speaking rate and memory can never be completely disregarded. By this account, it is assumed that speaking rate has a causal influence on phonological memory and not vice versa. The magnitude of the relationship between speaking rate and digit spans in the NH and CI groups and within the CI group should also not be disregarded. In comparing the relationship between speaking rate and digit span in these groups, several variations in the magnitude of the relationship were observed. These variations could be useful in determining why there are differences in the digit spans of the CI and NH children and the OC and TC children.

It is particularly interesting that the relationship between speaking rate and backward digit spans was observed in the CI group but was absent in the NH group. This finding suggests that deaf children with CIs may have used a different coding strategy to carry out this immediate memory task. The strong correlation of speaking rate with backward digit span in the CI users suggests that these children are using rehearsal strategies similar to what they are using in the digit span forward condition to complete the task. This strategy may, in fact, not be as efficient as the coding strategies that NH children typically use in tasks such as this. The differences in the correlations between backward digit spans and articulation rates in the CI and NH groups warrant further consideration. Additional research on executive functions of memory, such as organizing and planning recall mechanisms, are needed in order to determine why such differences are observed in the relationship between backward digit span and speaking rate in deaf children using CIs.

The differences in the correlations between speaking rate and digit span in the OC and TC groups are also important. Both the OC and TC groups showed robust correlations between the speaking rate and backward digit spans. However, the correlation between speaking rate and forward digit spans failed to reach significance in the OC group. In contrast, the correlation between speaking rate and forward digit span was particularly strong in the TC group. This discrepancy may be related to the differences observed in the two groups' forward digit span scores. The longer sentence durations observed in the TC group and the negative correlation with forward digit span suggest that slower speaking rates are more detrimental to digit span recall in the TC children than in the OC children. The TC children's speaking rates were also significantly slower than the OC children's speaking rates at all sentence lengths, a finding that provides additional support for this proposal.

The particularly strong relationship between speaking rate and forward digit span in the TC group suggests that the TC children's decreased rates of speaking and, therefore, rehearsing may adversely affect their phonological memory abilities relative to both the OC and NH children. In addition, the forward digit-span scores and span limits of the TC group appear to be closely tied to their response latencies. This relationship may also play a role in the shorter digit spans observed in this group. As the response latencies or *preparatory intervals* (Cowan et al., 1994) in the TC group increased, digit span scores and the maximum span decreased. This pattern indicates that the TC group may be receiving the least amount of benefit from the preparatory interval.

The preparatory interval is the final period of time in which items in the digit span lists can be actively rehearsed or refreshed in their entirety before response output. Instead of being assisted by this final opportunity, it appears that the TC users may be adversely affected by rapid memory decay that takes place during this time. This disadvantage may be related to the TC children's slower speaking rates. Because the TC users are unable to articulate quickly, they may be unable to subvocally rehearse fast enough during response preparation to counteract memory decay.

The OC group did not exhibit the same relationship between response latencies and digit spans. This finding indicates that their subvocal rehearsal rates may be fast enough to avoid the memory decay that could be hindering the TC group during response preparation. Although faster overt articulation likely contributes to the digit span difference observed within the CI group, faster overt articulation by the OC users was not directly related to their forward digit spans. This finding suggests that OC users may have an additional strategy or ability that benefits forward digit span recall. This additional processing strategy used in a phonological recall task may facilitate the longer forward digit spans observed in OC children.

One additional advantage that OC users may have over their TC counterparts is related to their memory scanning abilities exhibited during digit span recall pauses. Although not statistically significant, the interword pause durations observed in the OC children were consistently shorter than those found for the TC group. The shorter pause times indicate that the OC children were scanning items in short-term memory faster than the TC children. Faster scanning by OC children may have reduced the time during recall in which items could be forgotten, therefore facilitating their immediate memory in comparison to the TC children.

Clearly, the NH children had the advantage of faster scanning abilities during digit span recall. The NH children's interword pauses during recall were much shorter than those of the CI users. This suggests that the NH children were much more efficient at memory scanning. The differences in interword pause time measured in the NH and CI children's digit span responses may be partly responsible for the digit span differences observed between the groups. Faster scanning rates and the ability to speak and rehearse at faster speeds may be the primary factors responsible for the digit span differences observed between the NH and CI groups.

The overall pattern of results found in both the CI and NH groups is quite similar to those reported recently by Cowan et al. (1998) in NH children. Both studies suggest that covert verbal rehearsal and serial scanning speed of short-term memory processes are important factors affecting immediate memory span in NH children. Cowan et al. found that those children who were fastest at the two processes had longer memory spans. However, this study was restricted to NH children that differed only in chronological age.

Comparable results were observed in the present study using children with similar chronological ages but quite different developmental histories relating to early auditory experience. The similarities of the results between this study and Cowan's studies with NH children indicate that speed of articulation, rehearsal, and memory scanning (i.e., retrieval from short term memory) may be closely linked to early auditory experiences and activities involving speech and spoken language processing. The contribution of early auditory and speech experience found in this study suggests that subvocal rehearsal and serial scanning processes may not be exclusively related to developmental milestones that are more cognitively or metacognitively centered, such as the ability to effectively organize and utilize the two processes in tasks requiring immediate recall. Rather, efficient subvocal rehearsal and scanning may be strongly dependent on underlying mechanisms of auditory perception and speech production that contribute to the development of phonological processing skills and the active use of verbal rehearsal strategies in working memory.

Because the group of deaf children were found to fall within the normal range of intelligence prior to being recruited for the current study, the most probable developmental influence on their decreased speaking rehearsal speed, scanning rates and shorter digit spans is their early period of sensory deprivation. The majority of the CI users in this study were congenitally deaf or deafened shortly after birth, causing their early developmental experiences to be quite different from the NH children's experiences. The most obvious and influential difference in the developmental experiences of normal-hearing and deaf children is the sensory deprivation endured by deaf children. This sensory deprivation likely results in widespread developmental brain plasticity, further differentiating deaf children's development from that of normal-hearing children. Such brain plasticity affects the central auditory system, as well as other cortical areas, both before and after cochlear implantation (Ryugo, Limb & Redd, 2000). The breadth of cortical plasticity that is possible during a period of sensory deprivation may also play a role in the CI children's atypical performance on cognitive tasks such as digit span recall. In addition, it should be emphasized here that cochlear implantation itself does not remediate the hearing of deaf children to normal. Rather, children with cochlear implants must learn to use an altered electrical signal to perceive and produce speech (Balkany, Hodges, Miyamoto, Gibbin & Odabasi, 2001; Miyamoto & Kirk, 1999). This unique form of auditory perception is also a hallmark difference in the developmental course experienced by deaf children after cochlear implantation.

Taken together, all previously mentioned factors prevent profoundly deaf children with CIs from simply initiating a delayed "normal" course of auditory, speech, and language acquisition. Instead, deaf children with CIs may follow a somewhat different developmental pattern of speech and language acquisition that may play a major role in the speed at which speech is perceived and produced and how effectively it is subvocally represented, rehearsed and retrieved. Such differences are likely the primary influences contributing to the differences in immediate memory span that were observed in this study. These differences may also propagate and cascade up the information processing system to affect other cognitive domains, such as reading, learning and allocating attention to surroundings. For this reason, these domains should be included in future investigations of the perceptual and cognitive development of profoundly deaf children using cochlear implants.

Summary and Conclusions

In summary, an examination of articulation rates obtained from a sentence repetition task revealed that deaf children with CIs are much slower at producing overt speech than their NH peers. Additionally, measures made on the verbal recall portion of the WISC-III auditory digit span task showed that interword pauses taken during digit span recall are longer in deaf children using CIs than in NH children. Both sentence durations and interword pause durations displayed during recall were also longer in TC children than in OC children. The differences observed between the TC and OC children demonstrate the robust effects of early oral-aural experiences on covert and overt verbal rehearsal. The decreased rehearsal rate and retrieval speed from short-term memory are likely contributing to the shorter digit spans observed in the TC group. The atypical working memory performance observed in the deaf children with CIs suggests that their slower speeds of overt articulation result in slower covert verbal rehearsal processes which are crucial for maintaining information in short-term memory. Additionally, the longer pauses taken during recall reflect slower serial scanning and retrieval in the CI users, particularly those children who are placed in total communication environments.

The differences observed in verbal rehearsal and scanning speed in deaf children with CIs replicate recent findings in NH children who were studied at different points in development (Cowan et al., 1992; Cowan et al., 1994; Cowan et al., 1998; Cowan, 1999). Contrary to being broad developmental influences related to chronological age and maturation, the longer articulation times and slower scanning speeds observed in the present study with deaf children who use CIs may be the consequences of different early auditory and linguistic experiences caused by a period of auditory deprivation as well as the unusual electrical stimulation provided by a cochlear implant. The unique developmental experience of auditory deprivation prompting the intervention of cochlear implantation is the most likely influence preventing these children from developing the same proficient and highly automatic skills to process, produce, rehearse and retrieve verbal information from working memory that their NH peers have developed over the same time period.

References

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

Baddeley, A. (1992). Working Memory. *Science*, 255, 556-559.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review, 105*, 158-173.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36*, 233-252.

Baddeley, A., Thomson, & Buchanan. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 575-589.

Balkany, T., Hodges, A., Miyamoto, R., Gibbin, K., & Odabasi, Onur. (2001). Cochlear Implants in Children. *Otolaryngologic Clinics of North America, 34*, 455-467.

Banks, J., Gray, D., & Fyfe, R. (1990). The written recall of primed stories by severely deaf children. *British Journal of Educational Psychology, 60*, 192-206.

Bebko, J. (1984). Memory and rehearsal characteristics of profoundly deaf children. *Journal of Experimental Child Psychology, 38*, 415-428.

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.

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

Campbell, R., & Wright, H. (1990). Deafness and immediate memory for pictures: Dissociations between "inner speech" and "inner ear". *Journal of Experimental Child Psychology, 50*, 259-286.

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

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology, 87*, 193-208.

Cleary, M., Pisoni, D., & Geers, A. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear & Hearing, 22*, 395-411.

Conners, F. A., Carr, M.D., & Willis, S. (1998). Is the phonological loop responsible for intelligence-related differences in forward digit span? *American Journal on Mental Retardation, 103*, 1-11.

Cowan, N. (1992). Verbal memory and the timing of spoken recall. *Journal of Memory and Language*, *31*, 668-684.

Cowan, N. (1994). Mechanisms of verbal short-term memory. *Current Directions in Psychological Science*, *3*, 185-189.

Cowan, N. (1996). Short-term memory, working memory, and their importance in language processing. *Topics in Language Disorders*, *17*, 1-18.

Cowan, N. (1999). The differential maturation of two processing rates related to digit span. *Journal of Experimental Child Psychology, 72*, 193-209.

Cowan, N., Keller, T., Hulme, C., Roodenrys, S., McDougall, S., & Rack, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory and Language, 33*, 234-250.

Cowan, N., Wood, N., Wood, P., Keller, T., Nugent, L., & Keller, C. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology, 127*, 141-160.

Daneman, M., Nemeth, S., Stainton, M., & Huelsmann, K. (1995). Working memory as a predictor of reading achievement in orally educated hearing-impaired children. *The Volta Review, 97*, 225-241.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test, Third Edition*. Circle Pines, MN: American Guidance Service.

Elliot, J. (1992). Forward digit span and articulation speed for Malay, English, and two Chinese dialects. *Perceptual and Motor Skills*, *74*, 291-295.

Ellis, N., & Anders, T. (1969). Short-term memory in the mental retardate. *Journal of Experimental Psychology, 79*, 568-569.

Ellis, N., & Hennelly, R. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71*, 43-51.

Ferguson, A., Bowey, J., & Tilley, A. (2002). The association between auditory memory span and speech rate in children from kindergarten to sixth grade. *Journal of Experimental Child Psychology, 81*, 141-156.

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

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

Gathercole, S., Willis, C., Baddeley, A. & Emslie, H. (1994). The children's test of nonword repetitions: A test of phonological working memory. *Memory, 2,* 103-127.

Geers, A. (2000). Center for Childhood Deafness and Adult Aural Rehabilitation, Current research projects: Cochlear implants and education of the deaf child, third-year results. In *Central Institute for the Deaf Research Periodic Progress Report No. 35* (pp. 5-20). St. Louis, MO: Central Institute for the Deaf.

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development, 61*, 653-663.

Henry, L. (1994). The relationship between speech rate and memory in children. *International Journal of Behavioural Development, 17*, 37-56.

Hernandez, L. (1995). Current computer facilities in the Speech Research Laboratory. In *Research on Spoken Language Processing Progress Report No. 20* (pp.389-393). Bloomington, IN: Speech Research Laboratory, Indiana University.

Hitch, G., Halliday, M., & Littler, J. (1989). Item identification and rehearsal rate as predictors of memory span in children. *Quarterly Journal of Experimental Psychology, 41*, 321-337.

Hitch, G., Halliday, M., & Littler, J. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C., Maughan, S., & Brown, A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47*, 72-87.

Kail, R. (1997). Phonological Skill and Articulation Time independently contribute to the development of memory span. *Journal of Experimental Child Psychology*, *67*, 57-68.

Kail, R., & Hall, L. (1994). Processing Speed, Naming speed, and reading. *Developmental Psychology*, *30,* 949-954.

Kail, R., & Park, Y. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology, 57*, 281-291.

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

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

Leder, S., Spitzer, J., Kirchner, C., Flevaris-Phillips, C., Milner, P., & Richardson, F. (1987). Speaking rate of adventitiously deaf male cochlear implant candidates. *Journal of the Acoustical Society of America, 82*, 843-846.

Leonard, L. (1998). *Children with specific language impairment*. MIT Press. Cambridge, MA.

Liben, L., & Drury, A. (1977). Short-term memory in deaf and hearing children in relation to stimulus characteristics. *Journal of Experimental Child Psychology, 24*, 60-73.

Mackworth, J. (1964). Auditory Short-term memory. *Canadian Journal of Psychology*, *18,* 292-303.

McGarr, N. (1981). The Effect of Context on the Intelligibility of Hearing and Deaf Children's Speech. *Language and Speech*, *24,* 255-263.

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

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

Miyamoto, R., & Kirk, K. (1999). Controversies in cochlear implantation: Technical and surgical considerations (pp. 329-340). In R. Cotton, & C. Myer (Eds.), *Practical Pediatric Otolaryngology*. Philadelphia: Lippincott-Raven Publishers.

Murray, D., & Roberts, B. (1968). Visual and auditory presentation, presentation rate, and short-term memory in children. *British Journal of Psychology*, *59,* 119-125.

Naveh-Benjamin, M., & Ayres, T. (1986). Digit span, reading rate and linguistic relativity. *The Quarterly Journal of Experimental Psychology*, *38,* 739-751.

Nickerson, R. (1975). Characteristics of the speech of deaf persons. *Volta Review, 77*, 342-362.

Osberger, M., Maso, M., & Sam, L. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. *Journal of Speech and Hearing Research, 36*, 186-203.

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

Pisoni, D., & Geers, A. (1998). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language. *Research on Spoken Language Processing Progress Report No. 22* (pp. 336-343). Bloomington, IN: Speech Research Laboratory.

Pisoni, D. (1999). Individual differences in effectiveness of cochlear implants in prelingually deaf children: Some new process measures of performance. *Research on Spoken Language Processing Progress Report No. 23 Bloomington, IN: Speech Research Laboratory*, (pp. 1-47). Bloomington, IN: Speech Research Laboratory.

Pisoni, D., Cleary, M., Geers, A., & Tobey, E. (2000). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *The Volta Review, 101*, 111-164.

Powell, D., & Hiatt, M. (1996). Auditory and visual recall of forward and backward digit spans. *Perceptual and Motor Skills*, *82,* 1099-1103.

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

Ryugo, D., Limb, C., Redd, E. (2000). Brain Plasticity: The impact of the environment on the brain as it relates to hearing and deafness. In J. Niparko (Ed.), *Cochlear Implants, Principles and Practices* (pp. 33-56). Philadelphia: Lippincott Williams & Wilkins.

Standing, L., Bond, B., Smith, P., & Isely, C. (1980). Is the immediate memory span determined by subvocalization rate? *British Journal of Psychology, 71*, 525-539.

Sussman, J. (1993). Perception of formant transition cues to place of articulation in children with language impairment. *Journal of Speech and Hearing Research, 36*, 1286-1299.

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

Waters, G., & Doehring, D. (1990). Reading acquisition in congenitally deaf children who communicate orally: Insights from an analysis of component reading, language, and memory skills. In C. B. A. Levy (Ed.), *Reading and its development*. New York: Academic Press.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children, Third Edition (WISC-III)*. San Antonio, TX: The Psychological Corporation.

Williams, J. (1984). Phonemic analysis and how it relates to reading. *Journal of Learning Disabilities, 17*, 240-245.
Zifcak, M. (1981). Phonological awareness and reading acquisition. *Contemporary Educational Psychology, 6*, 117-126.

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

Perception of "Elliptical Speech" Following Cochlear Implantation: Use of Broad Phonetic Categories in Speech Perception[1]

Steven Fountain[2] and David B. Pisoni[2]

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

[1] This work was supported by NIH T32 Training Grant DC00012 to Indiana University. We thank Luis Hernandez for computer programming and support in this experiment, Rebecca Herman who voiced half and created all the stimuli used and Michael Vitevitch who voiced half of the stimuli used in this experiment. We also thank the ENT Hearing and Speech Clinic for providing the cochlear implant patient list and the setting where the experiment took place.
[2] DeVault Otologic Research Laboratory, Department of Otolaryngology, Indiana University School of Medicine, Indianapolis, IN.

Perception of "Elliptical Speech" Following Cochlear Implantation: Use of Broad Phonetic Categories in Speech Perception

Abstract. The present study investigated the perception of "elliptical speech" (Miller & Nicely, 1955) in 10 hearing-impaired adults with cochlear implants. A group of 15 normal-hearing adults were also tested for comparison. Sentence discrimination and repetition tasks were employed using sets of meaningful and anomalous English sentences. Two different versions of each set of sentences were constructed. One version contained sentences with intact place of articulation cues; the other version transformed the sentences into "elliptical speech" using a procedure in which the consonants were all replaced with other consonants that had the same voicing and manner of articulation features but always had an alveolar place of articulation. The hearing-impaired listeners who use cochlear implants completed a same-different sentence discrimination task and a repetition task under quiet listening conditions. The normal-hearing listeners completed both tasks under noise-masking conditions. In the same-different discrimination task, normal-hearing listeners perceived sentences with intact place of articulation cues and its elliptical version as the "same." The cochlear implant users showed a similar response pattern. These findings support Miller and Nicely's claim that under signal degradation conditions, ellipsis can no longer be detected reliably. In the repetition task, however, normal-hearing subjects showed somewhat better repetition performance for sentences with intact place of articulation than for elliptical speech sentences. This was unexpected given the earlier findings from the sentence discrimination task. The cochlear implant users also showed slightly better repetition performance for sentences with intact place of articulation cues than for elliptical speech. Taken together, both sets of findings on the perception of elliptical speech provide support for the hypothesis that hearing-impaired patients with cochlear implants perceive speech and recognize spoken words using broad perceptual equivalence classes. Even without highly detailed acoustic-phonetic information about place of articulation in speech, many patients with cochlear implants are able to reliably recognize words in sentences and repeat them immediately on-the-fly after only one presentation.

Introduction

What does speech sound like to a deaf patient with a cochlear implant? This is an important unexplored question in speech perception and spoken word recognition that has both theoretical and clinical implications. Do patients hear speech as a sequence of meaningful spoken words arrayed in time or do they perceive and recognize the degraded and impoverished nature of the speech signals that are transmitted by their cochlear implant? Clinical experience suggests that listeners who use cochlear implants often do not perform well on open-set tests of word recognition that require them to identify spoken words by accessing all possible words in their entire mental lexicon. Many of the confusions shown by deaf listeners who use cochlear implant in these word recognition tests are due to the problems in perceiving the place of articulation of consonants. However, despite this apparent difficulty in correctly perceiving place of articulation, many cochlear implant users are able to do very well in face-to-face conversation and other tasks that require comprehension of the intended message.

In their pioneering work on speech perception, Miller and Nicely (1955) found that place of articulation was frequently confused under masking and low-pass filtering. Based on their finding that some consonants tended to be confused with each other, Miller and Nicely grouped consonants into perceptual equivalence classes that they assumed were functionally the same under these degraded listening conditions. For example, if [p t k] are mistaken for one another frequently under degraded listening conditions, these segments would be placed in a single equivalence class. Miller and Nicely further suggested that a unique form of phonetically degraded speech could be created in which a single member of each equivalence class replaced every other individual member of its equivalence class. They called this kind of speech “elliptical speech” because of the ellipsis, or omission, of place of articulation information. For example, if speech is produced in which every [p t k] was simply replaced by [t], the resulting speech sounds very strange when presented in the clear. However, when this speech is presented under degraded listening conditions where the confusions were originally produced, Miller and Nicely predicted that elliptical speech should actually be undetectable because the members of the equivalence classes are grouped precisely under the same signal degradation conditions. Miller and Nicely reported some informal support for their predictions, although they never carried out a formal experiment to support their hypothesis (Miller, 1956; Miller & Nicely, 1955).

Recently, Quillet, Wright and Pisoni (1998) reported the results of a same-different discrimination experiment that was designed to assess the predictions made by Miller and Nicely that elliptical speech would be undetectable under conditions of signal degradation. Using synthesized speech, young normal-hearing listeners were presented with pairs of sentences that were either lexically the same or lexically different. In one condition, both sentences in a lexically identical pair were intact and were presented without any ellipsis. In a second condition, both sentences in a pair were transformed into elliptical speech. Finally, in a third condition, one sentence in a lexically identical pair was presented intact while the other was an elliptical version. Quillet et al. found that when the stimuli were presented in the clear with no signal degradation, the listeners identified most of sentences in this third condition as “different.” However, under signal degradation conditions, the majority of the listeners identified a large number of the pairs in this condition as “same” indicating that the ellipsis of place of articulation could not be detected under these conditions.

In addition to replicating Miller and Nicely’s informal experiment on elliptical speech and providing support for their earlier predictions, Quillet et al. (1998) also noted several interesting parallels in speech perception between normal-hearing listeners under conditions of signal degradation and deaf listeners who use cochlear implants. Just as normal-hearing listeners show systematic confusions among different places of articulation under conditions of signal degradation, patients with cochlear implants also show similar confusions among places of articulation. Quillet et al. suggested that it might be possible to use “elliptical” speech as a research tool to investigate the perception of speech by patients with cochlear implants and to try to understand how they do so well with their implant even with highly impoverished input signals. If we find that cochlear implant users perceive “elliptical speech” as the same as intact speech, this result would provide support for the hypothesis that cochlear implant users employ broad perceptual equivalence classes to perceive place of articulation in speech perception.

The present investigation is an extension of a recent case study conducted by Herman and Pisoni (2000). They tested an exceptionally good cochlear implant patient (“Mr. S”) under conditions of elliptical speech and 20 normal-hearing listeners using degraded speech signals. The design of the present experiment was similar to the preliminary study carried out by Herman and Pisoni. In both studies, a same-different discrimination task was employed to determine if elliptical speech would be undetectable to listeners who use cochlear implants. The two earlier investigations differed in terms of how the stimuli were created. Herman and Pisoni created their stimuli using natural speech whereas Quillet et al. used synthesized speech.

In the present study, pairs of English sentences were presented to adult cochlear implant users who were asked to determine whether the two sentences were the “same” or “different.” As in Quillet et al., the crucial test case was the third condition in which two sentences were lexically identical, but one was intact while the other was transformed into elliptical speech. In this condition, under the hypothesis that cochlear implant users cannot detect place of articulation information, we predicted that our patients would label the two sentences as the "same." This pattern of results would suggest that consonants with the same manner and voicing features but different places of articulation form an equivalence class and that cochlear implant users recognize words in context using broad phonetic categories. This pattern of results would also provide support for the hypothesis that cochlear implant users perceive speech as a sequence of highly familiar words and do not normally detect fine acoustic-phonetic differences between spoken words in meaningful sentences.

Up to this point, we have focused on what speech might sound like to users of cochlear implants, and thus what obstacles might have to be overcome to recognize words successfully. A second question of interest to us concerns why some users of cochlear implants manage to perceive speech so well in face-to-face conversations despite the degraded auditory signals they receive through their implants. One explanation for their good performance in the real world is the observation that powerful structural constraints influence the sound patterns found in all languages (Shipman & Zue, 1982; Zue & Huttenlocher, 1983). For example, Zue and Huttenlocher (1983, p. 122) observed that the sound patterns of words in spoken languages are highly constrained not only by the inventory of individual speech sounds in a particular language but also by the allowable combinations of those sound units" (i.e., the phonotactic constraints). Shipman and Zue (1982) reported that an analysis of English that distinguishes only between consonants and vowels can prune a 20,000-word lexicon down to about 200 CV combinations.

Because strong structural constraints on sound patterns exist in language, a broad phonetic classification can serve to define a "cohort" or the set of possible candidate words having the same pattern. As Shipman and Zue showed in their computational research, these candidate sets may actually be quite small. They found that the average size for these equivalence classes for the 20,000-word lexicon was approximately 2 and the maximum size was approximately 2,000 (Zue & Huttenlocher, p. 122). Thus, even if a listener does not accurately perceive the exact place of articulation, he can still identify words successfully using broad equivalence classes if he can recognize at least the sequence of consonants and vowels in the pattern. Broad phonetic coding of the input signal may be sufficient to permit higher-level lexical processes to take over and support word recognition and spoken language comprehension despite degraded input signals.

While research has shown that phonotactic constraints exist in language, listeners may not necessarily use them in real-time speech perception. Quillet et al. (1998) investigated whether coarse coding of the speech signal could provide a rich and sufficient set of cues to allow normal-hearing listeners to understand meaningful sentences. In a transcription task using synthesized speech, young normal-hearing listeners were asked to transcribe key words from each sentence. The sentences had either intact place of articulation cues or were produced using elliptical speech. The sentences were presented in the clear or in white noise at 0 dB SNR, -5 dB SNR, and -10 dB SNR. Quillet et al. predicted that while speech with intact place of articulation should show decreased intelligibility under degraded conditions, elliptical speech should actually show the reverse pattern, that is, increased intelligibility as signal degradation increased. Quillet et al. found that speech with intact place of articulation cues did show decreases in transcription accuracy under degraded conditions whereas the elliptical speech showed improvements in transcription accuracy from the 0 dB SNR level to the -5 dB SNR level before dropping off at the -10 dB SNR level. Quillet et al. interpreted these findings as support for the proposal that normal-hearing listeners are able to make use of broad phonetic categories to identify spoken words in sentences under conditions of signal degradation.

In their case study of "Mr. S," Herman and Pisoni (2000) explored whether one excellent cochlear implant user utilized coarse coding and broad phonetic categories in perceiving sentences. They used a transcription task that was similar to the task used by Quillet et al. However, Herman and Pisoni used natural speech rather than synthetic speech. They predicted that normal-hearing listeners and their cochlear implant patient would transcribe elliptical speech at the same level of accuracy as intact speech.

Herman and Pisoni found that the normal-hearing listeners and their cochlear implant patient transcribed intact speech more accurately in than elliptical speech, a result that differed from Quillet et al.'s earlier findings. The pattern of results indicates that some conflicting phonetic cues to place of articulation were actually perceived in the stimuli. These cues created confusions for the listeners in recognizing the elliptical speech samples. To explain the discrepancy between the two studies, Herman and Pisoni suggested that the redundant speech cues present in the natural speech stimuli might have survived the signal degradation more than Quillet et al.'s synthetic speech, which was much less redundant than the natural speech samples. Also, Herman and Pisoni provided their listeners with the option of repeating the sentence up to five times before they had to record their responses. This may also have improved their performance under the degraded conditions. In Quillet et al., listeners only heard each test sentence once.

The main difference between the earlier study by Herman and Pisoni (2000) and the present investigation is the format of the discrimination task. In Herman and Pisoni, a transcription task was employed in which the participants simply wrote down "keywords" of a sentence presented to them over a loudspeaker. For each sentence, a response frame with blanks substituted for the keywords was provided for them to transcribe their answers. Additionally, the participants had the option of listening to the sentence up to five times in order to fill in each blank. The methodology was changed for this study to a repetition task. Listeners were asked to repeat out loud as much of the sentence as possible after listening to the sentence only once. This was thought to be a better way to measure of the immediate perception of the sentences. Unlike the earlier transcription task, the participants did not have access to the redundancy of the phonetic cues that may have been utilized by listeners who could hear the sentence repeatedly up to five times.

Although these differences may explain part of the discrepancy in the results between Herman and Pisoni and Quillet et al., natural speech was still used to create conditions that better resembled the real world. Because of the difficulty in carrying out an immediate repetition task, only cochlear implant users who scored above a minimum satisfactory level in a standard open-set word recognition test were selected for this study. Again, half of the sentences were intact and half were transformed into elliptical speech. If the coarse-coding hypothesis of spoken word recognition is correct, cochlear implant users should show the same performance on sentences with intact place of articulation cues as on sentences transformed into elliptical speech. This pattern would indicate that coarse coding was sufficient for successful word recognition to be carried out in spoken sentences.

Method

Participants

Ten severe-to-profoundly deaf patients who use cochlear implants were recruited based on their age, length of time using their implant, and most recent CNC list score from the patient charts at the Department of Otolaryngology, Indiana University School of Medicine. All of the patients were postlingually deafened and had acquired language normally before the onset of their hearing loss. Each participant with a cochlear implant was between the ages of 24 and 71 and had used their implant for at least one year. A standard clinical measurement of speech perception performance, their CNC score, represents the percentage of words that a hearing-impaired person correctly recognizes immediately following presentation of each word in an open-set format. A CNC score of 30% was the criterion used for participation in this study. Table 1 shows a summary of the demographic details of the patients who used cochlear implants.

In addition to the 10 patients, 15 normal-hearing listeners between the ages of 21 to 40 were also recruited using an email distribution list at the IUPUI campus. Both students and staff of the university responded and received a payment of 10 dollars for their time participating in the experiment. These listeners received the same conditions as the cochlear implant patients except that half of their trials were run under degraded listening conditions. None of the listeners reported any hearing or speech problems at the time of testing. All participants were native speakers of American English.

Cochlear implant user	Years with implant	CNC word score
1	2	62
2	6	50
3	2	86
4	5	30
5	3	46
6	1	77
7	3	50
8	5	74
9	1	38
10	10	64

Table 1. Age, number of years with implant, and CNC score of each cochlear implant user.

Stimulus Materials

The stimulus materials used in this study consisted of 96 Harvard sentences (IEEE, 1969). Each sentence contained five key words with declarative or imperative sentence structure, taken from lists 1-10 of Egan (1948). In addition to the *normal* Harvard sentences, a set of 96 *anomalous* sentences were created by substituting random words of the same lexical category (noun, verb, etc.) into the Harvard sentences of lists 11-20 (Egan, 1948). The anomalous sentences were developed to block normal top-down semantic processing. The new words were selected from lists 21-70 of the Harvard sentences (Egan, 1948). New sets of "elliptical" sentences were generated using these two types of sentences through a process of featural substitution that was similar to the original procedures developed by Miller and Nicely (1955). The stops, fricatives, and nasal consonants in each of the five key words were replaced with a new consonant that preserved the same manner and voicing features of the original consonant but changed the place feature to an alveolar place of articulation. For example, "See the plane in the blue sky" would be changed to "See the tlane in the dlue sky." Liquids /r l/ and glides /y w/ were excluded from the substitution process. Intact sentences and their elliptical versions are listed in the appendix along with the equivalence classes of consonants. This method of replacing consonants with alveolar consonants follows Miller and Nicely's original method of creating elliptical speech and differs from the procedures used earlier by Quillet et al. (1998).

A male and a female talker were used to generate the stimulus materials. Each talker read half of the sentences aloud. Before the recording session, both talkers practiced saying the test sentences several times. The speaker attempted to use the same intonation pattern in both the intact and elliptical versions of an utterance. Sentences were recorded using a head-mounted Shure Model SM98A microphone and a Sony TCD-D8 DAT recorder. The recordings were segmented into individual utterances, converted to a single channel, and downsampled to 22,050 Hz using CoolEdit™.

For the signal degradation, a masking noise was created using Gaussian noise applied to each sentence to create another set of stimuli. Noise was added at a -5 dB signal-to-noise ratio. Each noise-masked file was then saved as a separate digital speech file for use during presentation of the stimuli to the listeners.

Procedures

Sentence Discrimination Task. The 10 cochlear implant patients heard the test stimuli over an AV570 Advent loudspeaker. They were given four practice trials in which they could adjust the volume of the speaker to a comfortable listening level. A Visual Basic program running on a PC recorded subject responses and controlled the experiment. The experiment was self-paced to each subject. Each pair of sentences was presented only once. There was a one-second interval between the two sentences in each pair. Responses were entered using the computer mouse to click on a dialog box labeled "same" or "different" on the computer monitor. The participants were given explicit instructions to use the label "same" only for pairs of sentences that sounded exactly word-for-word and sound-for-sound identical. Each cochlear implant patient was presented with 64 sentences in two blocks of 32 trials each. They heard a block of normal sentences followed by a block of anomalous sentences. Half of the sentences in each block were spoken by a male speaker and half by a female speaker. Likewise, half of the sentences in each block were composed of intact speech cues and half consisted of elliptical speech sentences. A Visual Basic randomizer program on the PC was employed to randomize the order of the stimuli in each block.

Fifteen normal-hearing listeners followed the same procedures except that the volume of the speaker was set at a comfortable level of 80 dB. Also, the normal-hearing listeners did not receive the sentences in separate blocks but heard them in random order. Since normal-hearing listeners received half of the sentences under degraded listening conditions, 128 pairs of sentences, twice the number that cochlear implant users received, were presented to them. This procedure insured that the same number of sentences would be presented in the clear to both the cochlear implant patients and the normal-hearing listeners. Sentences presented in the clear and under degraded conditions were presented to normal-hearing listeners in a random order.

	Lexically Different	Lexically Identical
Both intact speech	IaIz	IaIa
Both elliptical speech	EaEz	EaEa
One intact, one elliptical	IaEz	IaEa
	EaIz	EaIa

Table 2. The different pairs of sentences used in the sentence discrimination task.

All participants received eight pairs of test sentences, as shown in Table 2. Pairs of sentences that were lexically identical are marked with two subscript "a's." Pairs of sentences that were lexically different are marked with a subscript "a" and a subscript "z." The sentences with intact place of articulation cues are referred to as "I" while the elliptical sentences are referred to as "E."

Sentence Repetition Task. The second part of the experiment employed an immediate repetition task. Subjects heard a sentence one time and were asked to immediately repeat back as much of the sentence as possible. Different sets of sentences were used in this task to avoid repetition and priming effects from earlier sentences used in the discrimination task. A Visual Basic program running on a PC controlled the experiment. This experiment was also self-paced. After repeating a sentence, the participant used a mouse to press a "next sentence" dialog box on the computer monitor to advance to the next trial. A microphone connected to a tape recorder recorded the subjects' vocal responses on each trial.

Ninety-six sentences were presented. Half of the sentences were normal sentences and half were anomalous sentences. Likewise, half of the sentences in each set contained intact speech cues and half contained elliptical speech cues; half were spoken by the female talker and half by the male talker. Normal-hearing participants heard half of their sentences in noise and half in the quiet. The patients with cochlear implants heard all of their sentences in the quiet.

Responses were scored using a very strict criterion. Responses were scored as correct if and only if they exactly matched the intended word. In the elliptical speech conditions, a word was scored as correct if it was produced as the original English word and incorrect if it was produced with any of the degraded speech cues that were actually heard (i.e., the elliptical version). For example, if the target word was "blue" and the elliptical version that was heard in the sentence was "dlue," a response of "dlue" would be scored as "incorrect." If the subject said "blue," it would be scored as "correct." A score for each sentence was computed based on the correct repetition of the three keywords in each sentence. A score of 0,1,2, or 3 was assigned if no keyword, one keyword, two keywords, or three keywords respectively were "correctly" repeated.

Results

Sentence Discrimination

Normal-hearing Listeners. The average number of "same" responses from the 15 normal-hearing listeners for both normal and anomalous sentences is shown in the two panels in Figure 1. Sentences presented in the clear are shown by light bars; sentences presented under degraded conditions are shown by dark bars. The different sentence pairings (i.e., IaIz, EaEz) are presented on the abscissa.

Normal-hearing listeners performed as expected on both normal and anomalous sentences. For the first four pairs of sentences in which the two sentences were lexically different, almost all of the pairs were labeled "different." For the next two sentence pair types, in which the two sentences were exactly identical, nearly all pairs were labeled "same." In the critical conditions, the last two sentence pair types, in which the two sentences were lexically the same although one sentence contained intact speech cues and the other sentence contained elliptical speech cues, normal-hearing listeners perceived most of them as "different" when they were presented in the clear. However, when the sentences were presented under degraded conditions, normal-hearing listeners could not discriminate between intact speech and elliptical speech versions of the same sentences and responded "same" on almost all of the trials.

The responses from of the normal-hearing listeners were subjected to a 2x2x4 analysis of variance (ANOVA). The first factor was "meaningfulness." The two levels were normal vs. anomalous sentences. The second factor was signal degradation. The two levels were sentences heard in clear and sentences heard under noise masking. The third factor was the pair type. Four levels for pair type are shown in the different cells in Table 2 and are plotted on the abscissa in Figure 1: IaIz and EaEz (two different sentences, either both have intact speech cues or both have elliptical speech cues), IaEz with EaIz (two different sentences, one has intact speech cues and one has elliptical speech cues), IaIa with EaEa (the same sentence twice, both have intact speech or both have elliptical speech cues), and the crucial test cases of IaEa with EaIa (the same sentence twice, one with intact speech cues and one with elliptical speech cues).

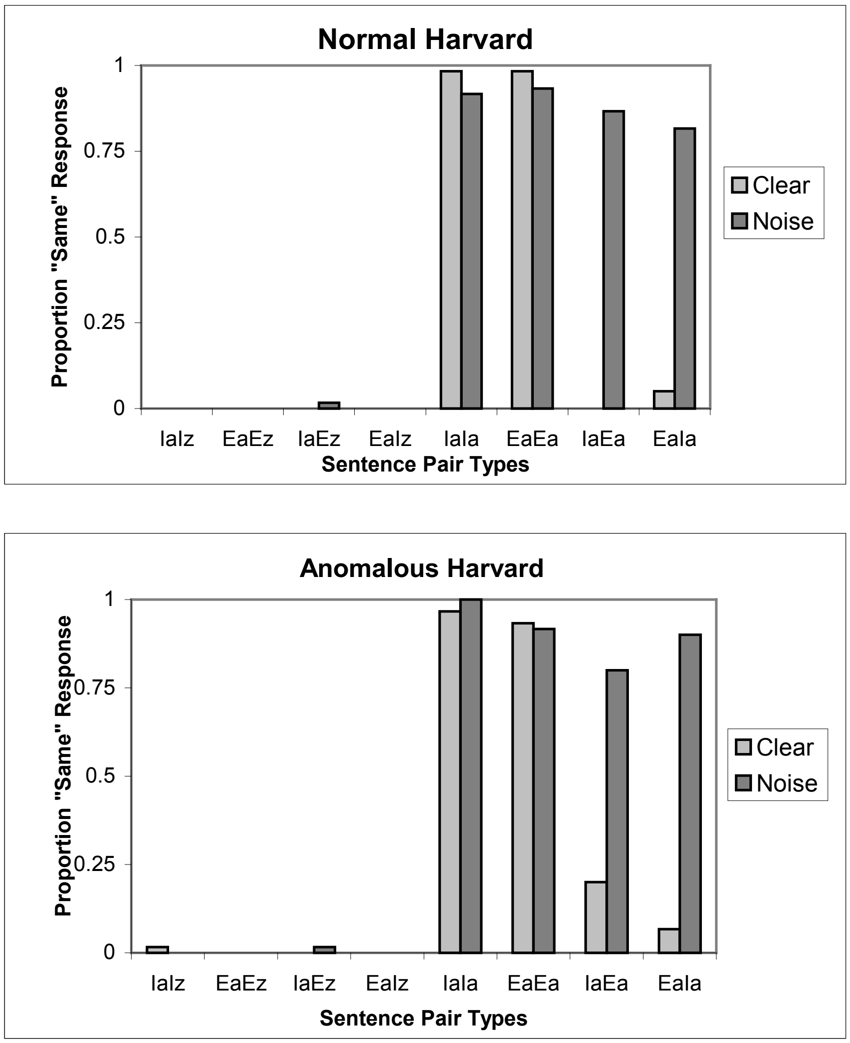


Figure 1. Results from the sentence discrimination task for normal-hearing listeners.

A significant difference in discrimination performance was obtained between sentences heard in the clear and sentences heard under degraded conditions (F (1,14) = 159.8, p <.001). The normal sentences and anomalous sentences did not differ significantly. The analysis also showed a significant two-way interaction between the signal degradation and pair type. Broken down by level using post-hoc tests for simple effects, signal degradation within the first level (IaIz and EaEz) and second level (IaEz and EaIz) showed no differences. When normal-hearing listeners heard two lexically different sentences, they had no difficulty with these pairs of sentences and were able to correctly discriminate these differences and respond appropriately regardless of whether the sentences contained both intact speech cues or both elliptical speech cues, or whether one sentence was intact and one was elliptical. Also, we found no difference in performance between sentences presented in the clear and sentences masked by noise. Within the third level (IaIa and EaEa) where identical sentences were presented twice, no difference was observed. Listeners correctly labeled the two sentences as the "same" most of the time. Again, we observed no difference in performance between the sentences presented in the clear and sentences masked by noise. These findings are consistent with our earlier predictions.

We did observe a significant difference for the "same" responses under signal degradation for the fourth set of sentences (IaEa and EaIa ($F(1,14) = 160.7$, $p < .001$)). For these critical pairs of sentences, two lexically identical sentences were presented. One sentence was intact and the other contained elliptical speech cues. In the clear, listeners correctly labeled almost all of these sentence pairs as "different," but under degraded listening conditions, almost all of the sentence pairs were labeled as "same." This is exactly the pattern of responding we expected to find if the normal-hearing listeners used broad equivalence classes to recognize spoken words in these sentences. And, this is precisely the pattern that would be predicted by Miller and Nicely based on the consonant confusions generated in noise.

Cochlear Implant Patients

A summary of the results for the 10 cochlear implant patients is shown in Figure 2. The results from the patients look quite similar to the data obtained from the normal-hearing listeners under degraded listening conditions. Lexically different sentence pairs were labeled "different" almost all of the time while lexically identical sentences were labeled "same" most of the time. The crucial test conditions were the last two sentences in which two lexically identical sentences were presented with one containing intact speech cues and the other containing elliptical speech cues. The results demonstrate that cochlear implant users for the most part could not discriminate differences between intact speech cues and elliptical speech cues. They perceived both types of elliptical sentences as the same more than 75% of the time.

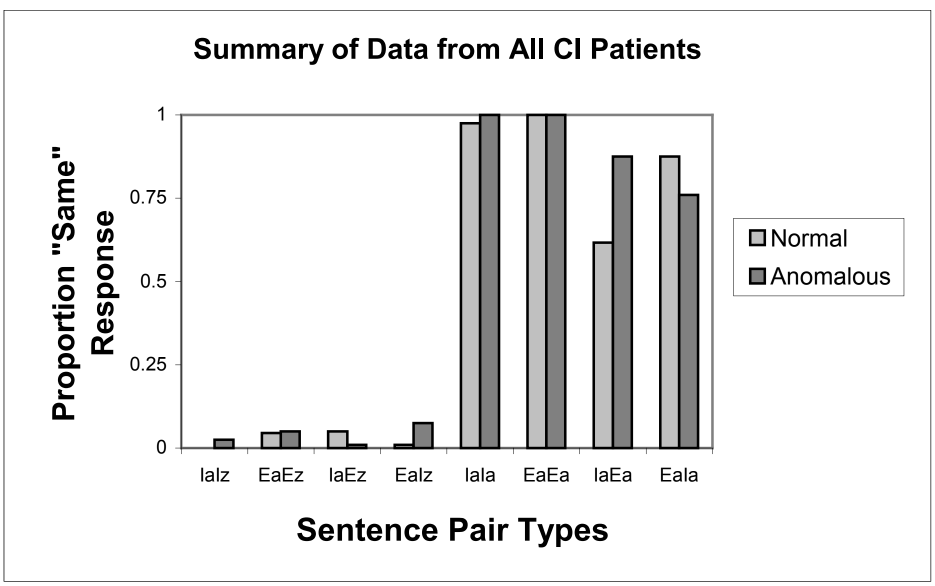


Figure 2. Mean results for the 10 cochlear implant patients.

The data were analyzed using a 2x4 ANOVA with the factors "meaningfulness" and "pair type" as in the normal-hearing listener's data analysis but without the additional factor of "degradation." No differences in discrimination were found between the normal and anomalous sentences.

To examine individual patient's data, three groups of listeners were created based on the earlier CNC word recognition scores. CNC scores from 30 to 40 were designated "fair" users, CNC scores from 41 to 62 were designated "good" users, and CNC scores from 63 to 86 were designated "excellent" users. Figure 3 shows individual data for the two fair users. Figure 4 shows individual data for the four good users. Figure 5 shows individual data for the four excellent users. Each patient's CNC score is shown next to his/her subject number in the legends of each figure. In each figure, normal sentences are shown in the top panel while anomalous sentences are shown in the bottom panel.

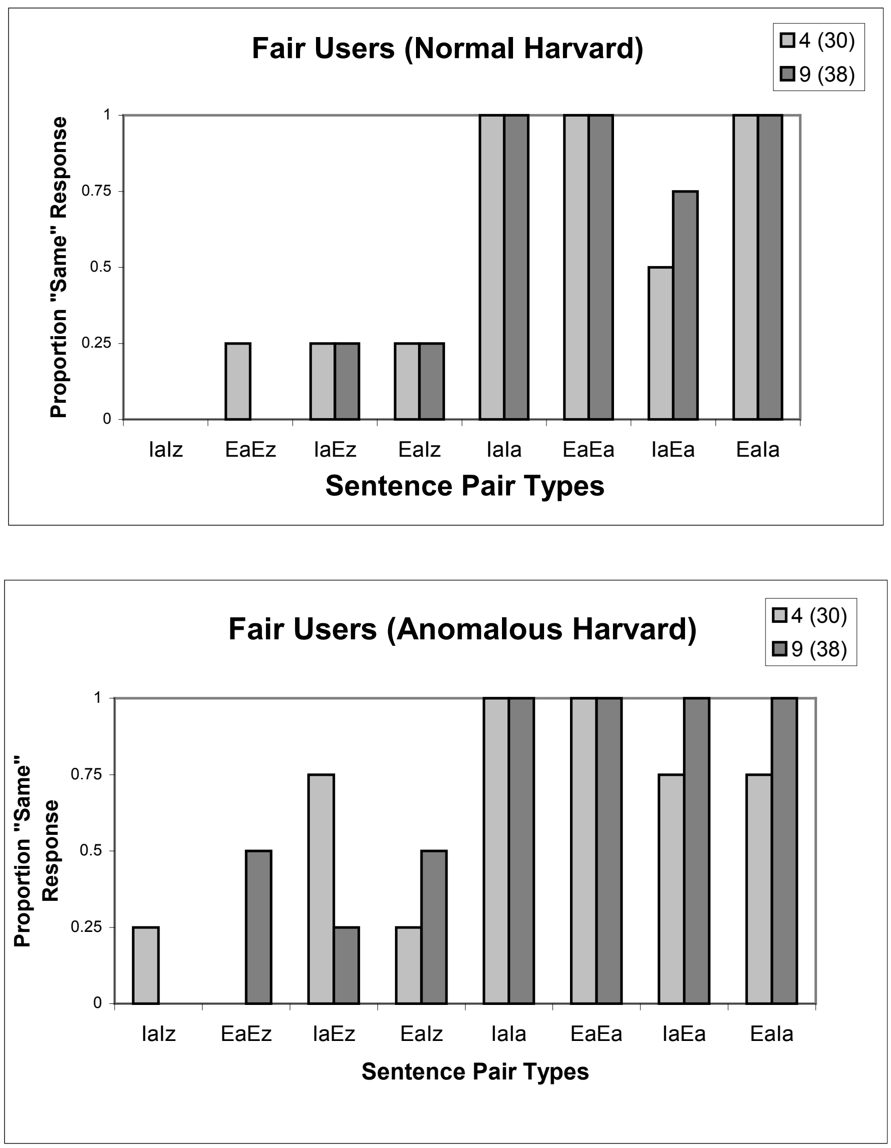
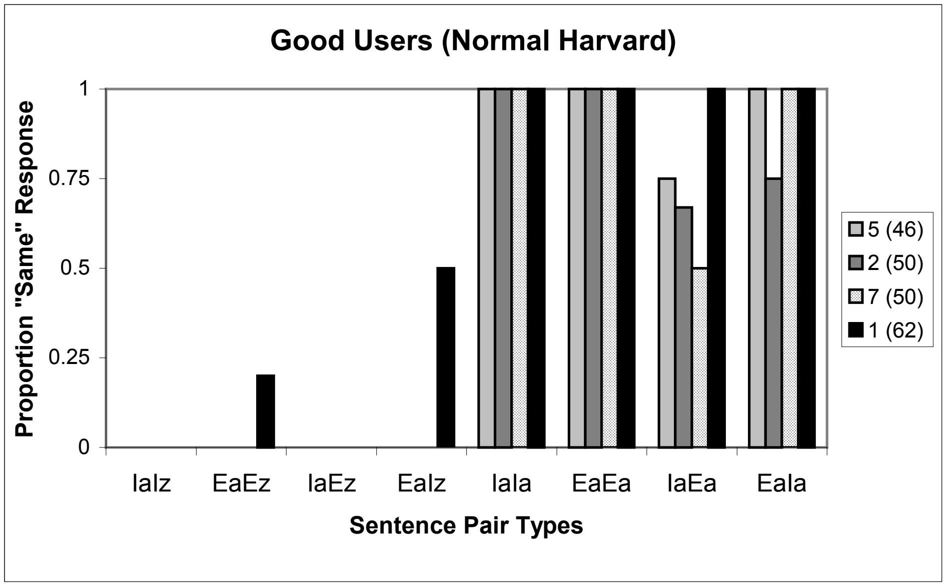


Figure 3. Individual results for the fair CI users.



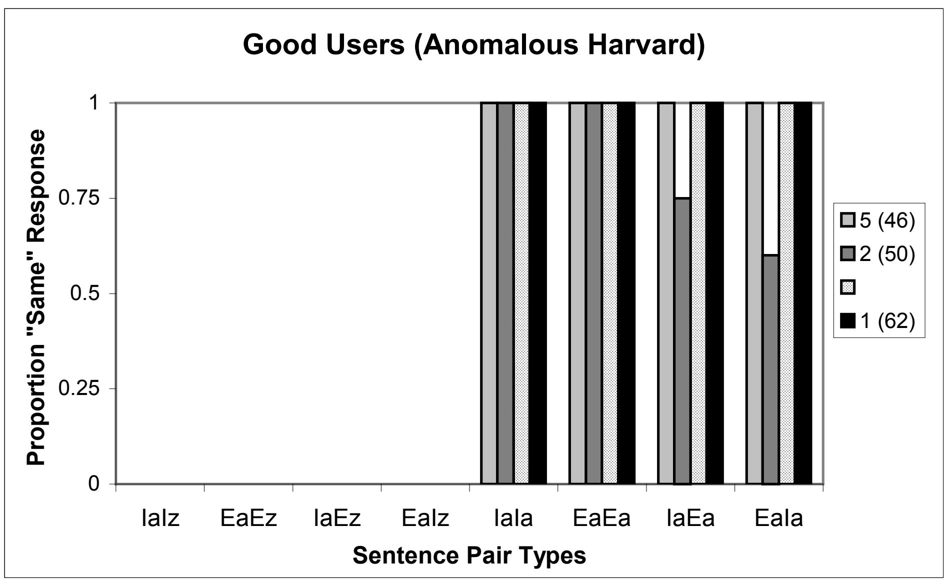


Figure 4. Individual results for the good CI users.

While the three groups of implant patients displayed response patterns that were similar to normal-hearing listeners under degraded listening conditions, the two extreme groups, the "fair" and "excellent" users showed somewhat different patterns of discrimination. The "fair users" had a great deal of difficulty distinguishing even between sentences that were lexically different. This is not surprising because they had low CNC scores to begin with. The "good users" rarely made that kind of error and the "excellent users" were always able to discriminate differences between lexically different sentences. On the other hand, the "excellent" users were able to discriminate differences between intact and elliptical speech and they performed better than the "good" users. The best user, Patient #3, with a CNC score of 86 was consistently able to discriminate between intact and elliptical speech cues in these sentences.

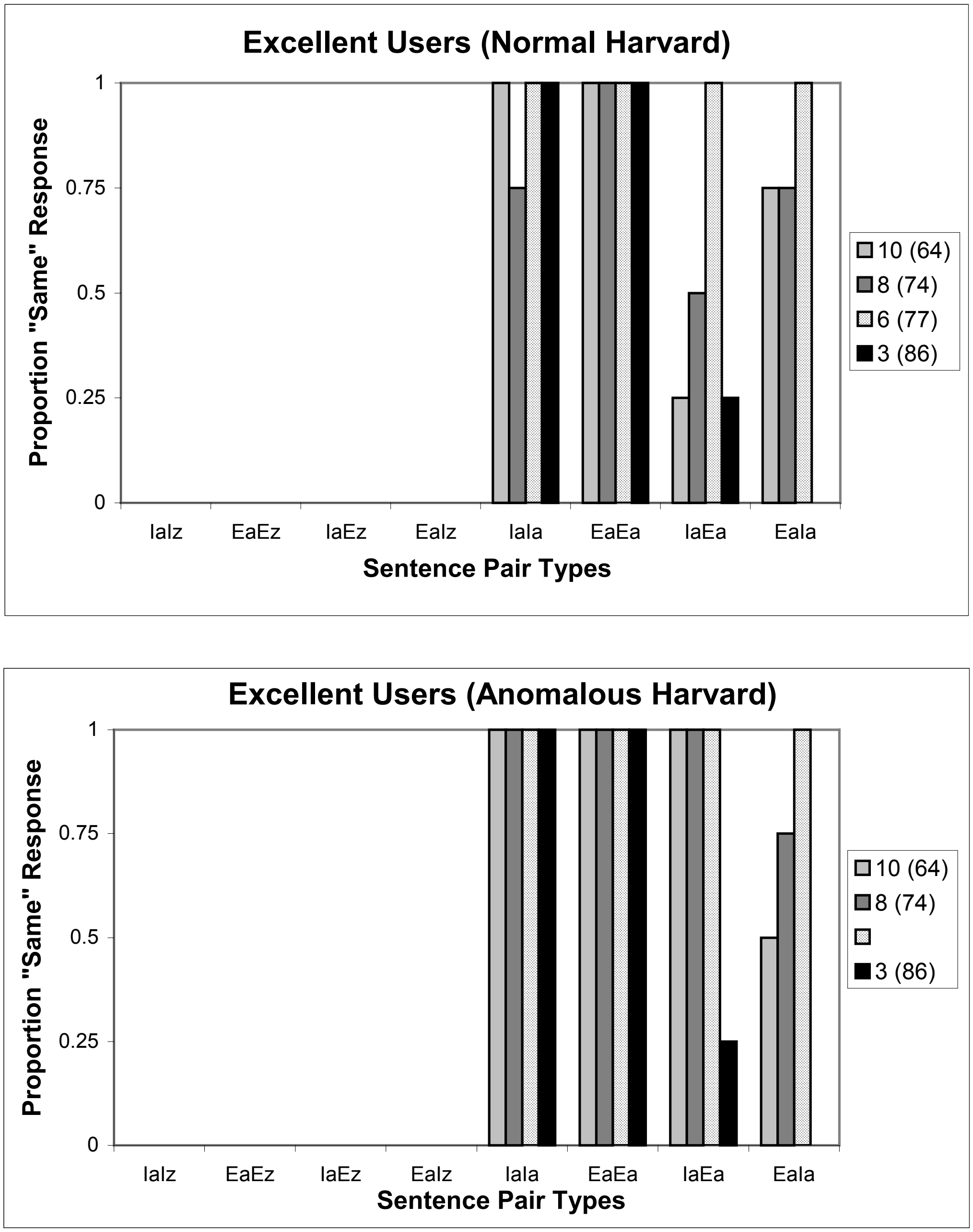


Figure 5. Individual results for the excellent CI users.

Sentence Repetition Task

Normal-hearing Listeners. The average correct proportion repetition performance for the 15 normal-hearing listeners for both normal and anomalous sentences is shown in Figure 6. Sentence Type is shown along the abscissa. The normal sentences are shown on the left; the anomalous sentences are shown on the right. Proportion correct indicates the proportion of intact word responses. Sentences presented in the clear are represented by the light bars and sentences presented under noise masking are shown by the dark bars.

The pattern of results is consistent with our earlier expectations. In the clear, normal-hearing listeners repeated words much better from intact sentences than elliptical sentences. This is due to our method of scoring since normal-hearing listeners repeated the words that were presented to them. Thus, in the clear, a sentence containing elliptical speech cues would be repeated with the incorrect speech cues present. However, for sentences in noise, normal-hearing listeners scored about the same for sentences containing intact speech cues and elliptical speech cues.

A 2x2x2 ANOVA was used to assess differences in the scores. The three factors were (a) speech cues (intact vs. elliptical), (b) signal degradation (clear vs. noise), and (c) meaningfulness (normal vs. anomalous sentences). Significant main effects were found for each factor: (a) $F\ (1,14) = 1889.6$, $p<.001$, (b) $F\ (1,14) = 1071.3$, $p<.001$, (c) $F\ (1,14) = 15.6$, $p = .001$, respectively.

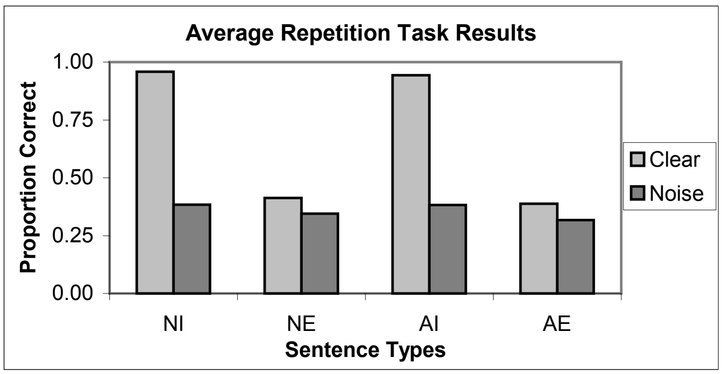


Figure 6. Average results from the repetition task for normal-hearing listeners. NI =normal Harvard intact sentence, NE =normal Harvard elliptical sentence, AI =anomalous Harvard intact sentence, AE =anomalous Harvard elliptical sentence.

For normal sentences, a separate 2x2 ANOVA revealed a significant main effect of speech cues ($F\ (1,14) = 1686.6$, $p<.001$). A significant main effect was also observed for signal degradation ($F\ (1,14) = 473.9$, $p<.001$). A post-hoc test for simple effects showed a significant difference between intact and elliptical speech cues for the normal sentences heard in noise ($F\ (1,14) = 8.38$, $p< 0.5$). For the anomalous sentences, a 2x2 ANOVA showed a significant difference for both speech cues ($F\ (1,14,) = 935.3$, $p<.001$) and signal degradation ($F\ (1,14) = 841.1$, p<.001). A post-hoc test revealed a significant difference ($F\ (1,14) = 44.6$, $p = .001$) between intact and elliptical speech of anomalous sentences heard in noise. Although normal-hearing listeners under degraded conditions appeared to score about the same for sentences containing intact speech cues and sentences containing elliptical speech cues, the analysis showed a significant difference between these two conditions. This difference was small compared to the difference in the scores when both sets of sentences were presented in the clear.

Cochlear Implant Patients. The mean sentence repetition performance results for six of the 10 cochlear implant users for both normal and anomalous sentences are shown in Figure 7. The remaining four cochlear implant users, whose CNC scores were very low, were unable to carry out the repetition task at all and their data were not included in this analysis.

The results shown in Figure 7 were surprising. The six cochlear implant users scored better on sentences containing intact speech cues than sentences containing elliptical speech cues. We expected that cochlear implant users would show performance that was similar to data obtained from normal-hearing listeners under signal degradation; however, the patients' performance was more similar to normal-hearing listeners in the clear. Somehow, the cochlear implant patients were able to distinguish between the two forms of speech.

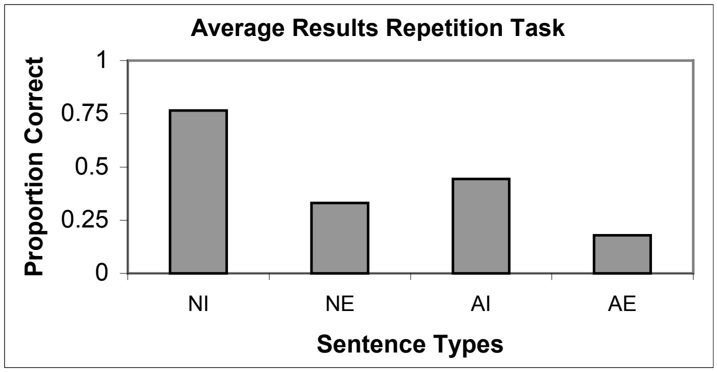


Figure 7. Average results from the repetition task for 6 cochlear implant patients.

The results were analyzed using a 2x2 ANOVA. One factor was speech cues (intact vs. elliptical cues). The second factor was meaningfulness (normal vs. anomalous sentences). Significant main effects were observed for both speech cues (F (1,5) = 150.1, p<.001) and sentence types (F (1,5) = 221.8, p<.001) suggesting that the cochlear implant patients were able to discriminate between intact and elliptical speech in this sentence repetition task.

The individual scores for the six cochlear implant patients that were able to complete the repetition task are shown in Figure 8 in order of increasing CNC score. An implant patient's CNC score is shown next to the corresponding number in the legends of each figure. Each of the patients scored better on sentences containing intact speech cues than the sentences containing elliptical speech cues. Performance on the sentence repetition task coincides with the CNC score for almost all of the cochlear implant users for each sentence type.

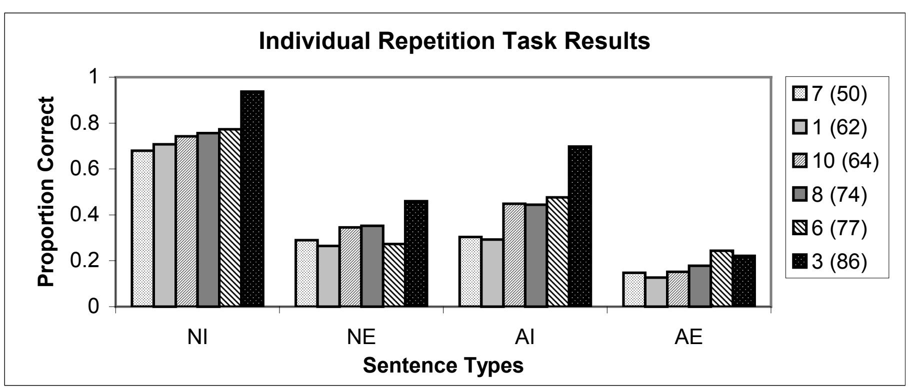


Figure 8. Individual results from the repetition task for the 6 cochlear implant patients.

General Discussion

Despite difficulties in perceiving fine phonetic contrasts in speech, such as place of articulation in consonants, many cochlear implant users are able to comprehend fluent speech and recover the talkers' intended linguistic message. What does the speech sound like for users of cochlear implants? How do patients with cochlear implants manage to comprehend spoken language despite receiving highly degraded sensory information? The results obtained from the two experiments reported in this paper provide some interesting new insights into the underlying perceptual processes and suggest some possibilities for intervention and oral rehabilitation in adult patients in the weeks and months immediately after they receive their cochlear implant.

Same-Different Discrimination Task

The first condition in this study used a same-different discrimination task with pairs of sentences that had either intact place of articulation cues or elliptical speech cues. The findings from normal-hearing listeners replicated the informal observations made by Miller and Nicely (1955) based on confusion data obtained in noise or under lo-pass filtering. Miller and Nicely originally suggested that elliptical speech, which is impoverished with respect to specifying place of articulation, may not be perceived as deficient under degraded listening conditions because these conditions "reinstate" or "reproduce" the original conditions that produced the signal degradation.

The present findings also replicate the earlier results reported by Quillet et al. (1998), which employed a similar same-different task and the results from Herman and Pisoni (2000) who used the same procedures with an exceptionally good patient with a cochlear implant. In both of these earlier experiments and the present study, normal-hearing listeners labeled pairs of sentences that were lexically identical but different because one contained intact speech cues and the other contained elliptical speech cues, as the "same" a majority of the time when heard under degraded conditions. When pairs of sentences were presented in the clear, normal-hearing listeners could easily discriminate intact speech from elliptical speech. However, when presented under degraded listening conditions such as masking noise, place of articulation information became less reliable and listeners tended to be unable to discriminate intact and elliptical speech. Listeners responded "same" almost all of the time based on broad phonetic coding of the stimulus input.

The cochlear implant patients that we studied here were also not able to distinguish intact speech cues from elliptical speech cues most of the time. This pattern was shown in Figure 2. The present results with a group of cochlear implant patients replicate the findings obtained in the case study reported by Herman and Pisoni (2000) who used only one cochlear implant patient. The present results suggest that phonetic contrasts such as place of articulation in consonants may not be perceived very well by cochlear implant users despite the fact that they can recognize spoken words and understand sentences. From an examination of the individual data, we observed a great deal of variability in performance across the 10 cochlear implant users. The "fair" users with CNC scores in the range from 30-40 had a great deal of trouble even discriminating sentences that were lexically different. In contrast, the best implant user who had a CNC score of 86 could consistently discriminate between intact and elliptical speech most of the time. The remaining users who fell somewhere between these two extremes appeared to produce results that were similar to normal-hearing listeners under degraded listening conditions.

The overall pattern of results observed in the same-different discrimination task by normal-hearing listeners under degraded listening conditions and by cochlear implant users in the quiet were similar to each other despite some small differences in procedure. Both groups appeared to rely on coarse phonetic coding in which place of articulation differences were no longer perceptually prominent. Both groups perceived the intact version and the elliptical versions of a sentence as the “same” under these presentation conditions. The present findings demonstrate the use of broad perceptual equivalence classes in speech perception when only partial stimulus information is present in the signal as originally suggested by Miller and Nicely (1955). This pattern of results resembles the earlier results reported by Herman and Pisoni (2000) and the preliminary findings from Quillet et al. (1998) for normal-hearing listeners under degraded listening conditions. Despite difficulty making fine lexical discriminations between spoken words, hearing-impaired individuals who use cochlear implants are able to use coarse-coding strategies to support spoken word recognition and speech perception. Many of these patients are able to derive substantial benefits from their cochlear implants even when they receive degraded and somewhat impoverished acoustic-phonetic information.

Sentence Repetition Task

The second condition in this study used an immediate repetition task with sentences that had either intact place of articulation cues or elliptical speech cues. These sentences were presented either in the clear or in noise. The findings from the normal-hearing listeners failed to support our initial predictions that repetition performance for intact speech and elliptical speech under degraded listening conditions would be equivalent. While the results for intact and elliptical speech in noise look very similar on average, statistical analysis revealed a small but significant difference between the two conditions. Thus, sentence repetition performance for normal-hearing listeners under degraded conditions for the intact speech was higher than performance of elliptical speech.

The results from the sentence repetition task were not consistent with Miller and Nicely’s (1955) predictions or the earlier findings of Quillet et al. (1998) using a transcription task. However, the results replicate the more recent findings reported by Herman and Pisoni (2000) who reported that overall transcription performance under degraded conditions was slightly better for intact speech than elliptical speech. Their results suggest that some phonetic cues to place of articulation can be extracted from these stimuli. Our findings suggest that some additional phonetic cues to place of articulation are present in natural speech and these are not eliminated by using elliptical speech.

The performance of the cochlear implant patients on normal and anomalous sentences with intact and elliptical speech cues presented in the clear and in noise followed the same basic pattern observed in the normal-hearing listeners under degraded listening conditions. The pattern of results observed with these listeners also failed to support our original predictions. Similar to the results from Herman and Pisoni’s single cochlear implant user, each cochlear implant user’s repetition performance, regardless of their CNC word score, showed worse repetition performance for elliptical speech relative to speech with intact place of articulation. As shown in the individual results, the recognition of both intact and elliptical speech increased with increases in the CNC word recognition scores, although both conditions were quite different. This pattern suggests that like normal-hearing listeners who were able to perceive speech under degraded conditions, cochlear implant users were also able to perceive and make use of some weak phonetic cues even with presumed loss of information about place of articulation.

Although failing to support our original predictions, the patients with cochlear implants did show good transcription performance for the normal sentences compared to anomalous sentences, despite the impoverished phonetic input from their implants. The earlier observations of Shipman and Zue (1982) and Zue and Huttenlocher (1983) concerning the powerful sound sequencing constraints in English and their role in spoken word recognition are consistent with the present findings obtained from cochlear implant users in both tasks. The patients are clearly able to make extremely good use of the minimal speech cues available to them in order to reduce the search space and to permit reliable lexical access and sentence comprehension. In many ways, these are impressive accomplishments given the highly degraded nature of the acoustic-phonetic input these patients receive.

The use of elliptical speech as a research tool to study sentence perception may lead not only to a better understanding of which speech sounds patients can discriminate after cochlear implantation and which ones they cannot, but these new stimulus materials may also help us to create better methods for training awareness of difficult phonological contrasts in patients with cochlear implants and help them deal with speech and spoken language in more efficient and optimal ways as they gain more experience with their cochlear implant and the nature of the transformation it imposes on speech signals.

In summary, the present investigation of normal-hearing listeners and a small group of cochlear implant patients provided an interesting pattern of results that differed from our original predictions. In the sentence discrimination task, both normal-hearing listeners and cochlear implant users were unable to discriminate between intact and elliptical speech suggesting the use of broad equivalent classes in speech perception and spoken word recognition. However, in the sentence repetition task, both groups were able to discriminate between intact and elliptical speech cues. This pattern of results suggests that some acoustic-phonetic cues to place of articulation can be perceived reliably. Although we observed differences in the two tasks, the present findings suggest that patients with cochlear implants perceive speech using broad perceptual equivalence classes and they appear to make use of less detailed phonetic categories than normal-hearing listeners routinely do under the same listening conditions.

References

Egan, J.P. (1948). Articulation testing methods. *Laryngoscope, 58,* 955-991.

Herman, R. & Pisoni, D.B. (2000). Perception of “elliptical speech” by an adult hearing-impaired listener with a cochlear implant: some preliminary findings on coarse-coding in speech perception. In *Research on Spoken Language Processing Progress Report No. 24* (pp. 87-112) Bloomington, IN: Speech Research Laboratory, Indiana University.

IEEE. (1969). IEEE recommended practice for speech quality measurements. *IEEE Report No. 297*.

Miller, G.A. (1956). The perception of speech. In M. Halle (Ed.) *For Roman Jakobson* (pp. 938-1062.) The Hague: Mouton.

Miller, G.A. & Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*, 338-352.

Quillet, C., Wright, R. & Pisoni, D.B. (1998). Perception of “place-degraded speech” by normal-hearing listeners: Some preliminary findings. In *Research on Spoken Language Processing Progress Report No. 22* (pp. 354-375). Bloomington, IN: Speech Research Laboratory, Indiana University.

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*.

Zue, V.W. & Huttenlocher, D.P. (1983). Computer recognition of isolated words from large vocabularies. *ICASSP 83: Proceedings, April 14-16, 1983, Sheraton-Boston Hotel, Boston, Massachusetts, IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. 121-125). New York: The Institute.

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

Speech Feature Discrimination in Deaf Children Following Cochlear Implantation[1]

Tonya R. Bergeson,[2] David B. Pisoni,[2] and Karen Iler Kirk[2]

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

[1] This work was supported by NIH-NIDCD Training Grant DC00012 to Indiana University and by NIH-NIDCD Research Grant DC00064 to the Indiana University School of Medicine.
[2] Also DeVault Otologic Research Laboratory, Department of Otolaryngology-Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, IN.

Speech Feature Discrimination in Deaf Children Following Cochlear Implantation

Abstract. Speech feature discrimination is a fundamental perceptual skill that is often assumed to underlie word recognition and sentence comprehension performance. To investigate the development of speech feature discrimination in deaf children with cochlear implants, we conducted a retrospective analysis of results from the *Minimal Pairs Test* (Robbins et al., 1988) selected from patients enrolled in a longitudinal study of speech perception and language development. The MP test uses a 2AFC procedure in which children hear a word and select one of two pictures (bat-pat). All children were prelingually deafened, received a cochlear implant before 6 years of age, and used either oral or total communication. Children were tested once every six months to a year for seven years; not all children were tested at each interval. By two years post-implant, the majority of these children achieved near-ceiling levels of discrimination performance for vowel height, vowel place, and consonant manner. Most of the children also achieved plateaus but did not reach ceiling performance for consonant place and voicing. Finally, children's performance was related to measures of speech feature discrimination, spoken word recognition, and sentence comprehension.

Introduction

Speech feature discrimination is assumed to be a fundamental perceptual skill that underlies word recognition and sentence comprehension performance in normal-hearing listeners. With the advent of cochlear implants in young children, it is important to understand how effectively deaf children use input from their cochlear implants to discriminate such speech features. Previous research has shown that deaf children with cochlear implants (2 years post-implant) perceive spoken words using "broader" phonetic categories than normal-hearing children normally do (Pisoni et al., 1999). However, do deaf children with cochlear implants eventually catch up with their normal-hearing peers in terms of discriminating phonetic features in words?

To investigate the development of speech feature discrimination in deaf children with cochlear implants, we conducted a retrospective analysis of results from the *Minimal Pairs Test* (Robbins et al., 1988) administered to patients enrolled in a longitudinal study of speech perception and language development at the Indiana University School of Medicine. A minimal pair consists of two words that differ in sound by only one feature (e.g., key-pea, big-bug). The MP test uses a 2-alternative forced-choice procedure in which children hear a spoken word and select one of two pictures that depict a minimal pair. If speech feature discrimination underlies word recognition and sentence comprehension performance, we would also expect the results from the *Minimal Pairs Test* to be correlated with measures of spoken word recognition and sentence comprehension. Thus, we also conducted a retrospective analysis of results from the *Phonetically Balanced Kindergarten Words* (PBK) test (Haskins, 1949), a test of spoken word recognition, and the *Common Phrases* test (Osberger et al., 1991), a test of sentence comprehension. In these tests, children are asked to repeat either a word or a sentence spoken by the clinician. All children in the present study were prelingually deafened, received a cochlear implant before 6 years of age, and used either oral or total communication. Children were tested once every 6 or 12 months for seven years. We predicted that the children's performance on all tests would improve over time and that children with oral communication experience would perform more accurately compared to children with total communication experience, particularly on the tests with an open-set response format. Moreover, we predicted that children's scores on tests of speech feature discrimination, spoken word recognition, and sentence comprehension would be positively correlated, consistent with the notion that speech feature discrimination skills underlie word recognition and sentence comprehension.

Method

Data Selection

We conducted a retrospective analysis of results from a subset of tests administered to patients enrolled in a longitudinal study currently being conducted at the Indiana University School of Medicine. Children were tested once every 6 or 12 months for seven years. Not all children were tested at each interval because of time constraints, child inattention, or cancelled appointments due to family issues.

Participant Characteristics

Table 1 shows a summary of participant characteristics. All children in the present study were prelingually deafened (i.e., deafened under the age of three years) and received a cochlear implant between one and six years of age. The age at the time of the first test session ranged from 2.8 to 6.9 years. Finally, children were tested over a span of three to seven years and accumulated at least four years of cochlear implant use by the date of the last test session. All tests were administered by speech-language pathologists or audiologists in a quiet room at Riley Hospital for Children, Indianapolis, IN.

	Age at Onset	Age at Implantation	Age at First Test
Oral Communication (N = 18)			
Average (years)	0.35	4.2	5.8
Minimum (years)	0	1.4	4.5
Maximum (years)	2.42	5.8	6.9
Total Communication (N = 18)			
Average (years)	0.36	4.1	5.0
Minimum (years)	0	2.2	2.8
Maximum (years)	1.67	5.8	6.7

Table 1. Oral Communication and Total Communication participant characteristics.

Dependent Measures

Speech Feature Discrimination. The *Minimal Pairs Test* (Robbins et al., 1988) is a closed-set test of speech feature discrimination, with a total of 80 minimal pair items. A minimal pair consists of two monosyllabic words that differ in sound by only one feature (e.g., key – pea, big – bug; see Figure 1). Target items were presented in a live-voice auditory-only format. That is, the clinician covered his or her face with a mesh screen while saying the target item aloud. The child was asked to point to only one of the two pictures corresponding to what they perceived.

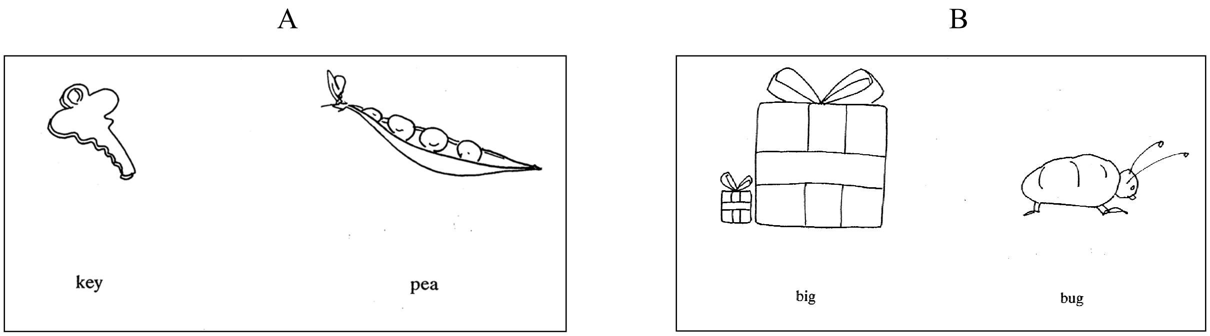


Figure 1. Examples of plates from the *Minimal Pairs Test* (Robbins et al., 1988). A) "Key" and "pea" differ by consonant place. B) "Big" and "bug" differ by vowel place.

Spoken Word Recognition. The *Phonetically Balanced Kindergarten Words* (PBK) test (Haskins, 1949) is an open-set test of spoken word recognition, with a total of 25 items in each list. There are four available lists of 50 words, but only three lists and 25 of the 50 words on each list were used in the present study. Each child was presented with only one list in each testing session. Like the *Minimal Pairs Test*, words in the PBK test were presented in an auditory-only live-voice format. Children were then asked to repeat the word they perceived. For example, the clinician said "sled" and the child should have responded, "sled" if they correctly heard the word. Children's responses were scored according to percent words correct.

Sentence Comprehension. The *Common Phrases* test (Osberger et al., 1991) is an open-set test of sentence comprehension, with a total of 30 items. Sentences were presented in one of three formats: Auditory-only (CPA), Visual-only (CPV), and Combined Auditory+Visual (CPAV). Children heard and/or saw a sentence spoken by the clinician and were asked to repeat the sentence back to the clinician or answer appropriately. For example, the clinician might say, "When is your birthday?" The child's response would be considered correct if he or she repeated, "When is your birthday?" or answers "August."

Independent Measures

Communication mode was determined by consulting the medical charts for each child on file at the Indiana University Medical Center.

Results

By 2 years post-implant, most children achieved near-ceiling levels of discrimination performance for vowel height, vowel place, and consonant manner (see Figures 2 & 3). Most children achieved plateaus in performance, but not ceiling performance, for consonant place and voicing. Consistent with previous findings in closed-set tests, there were no overwhelming differences in performance between Oral Communication and Total Communication children. However, Oral Communication children's performance was consistently more accurate across all tests compared to that of Total Communication children.

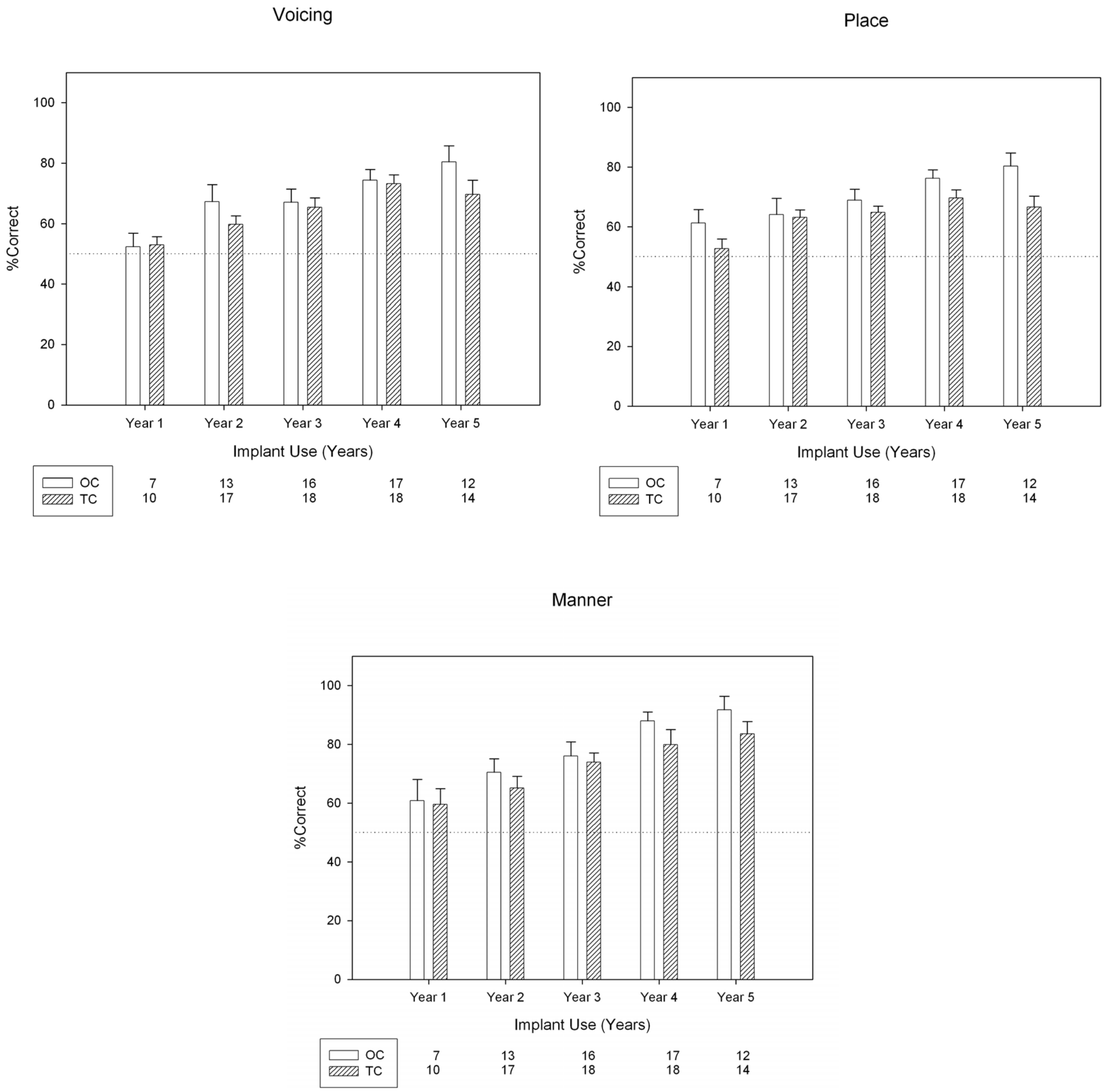


Figure 2. Percent correct for consonant features in the Minimal Pairs Test. Numbers below each graph represent sample size for the Oral Communication and Total Communication groups at each year of implant use. Error bars represent standard error.

Spoken word recognition and sentence comprehension scores increased over time but never reached ceiling levels, regardless of communication mode (see Figure 4). Nevertheless, Oral Communication children exhibited superior performance for measures of spoken word recognition and sentence comprehension compared to Total Communication children. Speech feature discrimination skills appear to increase in parallel with spoken word recognition and sentence comprehension skills.

Speech feature discrimination performance (Minimal Pairs) was strongly correlated with open-set spoken word recognition performance (PBK) and sentence comprehension (Common Phrases) (see Table 2). The correlations were stronger for Oral Communication children compared to those for Total Communication children.

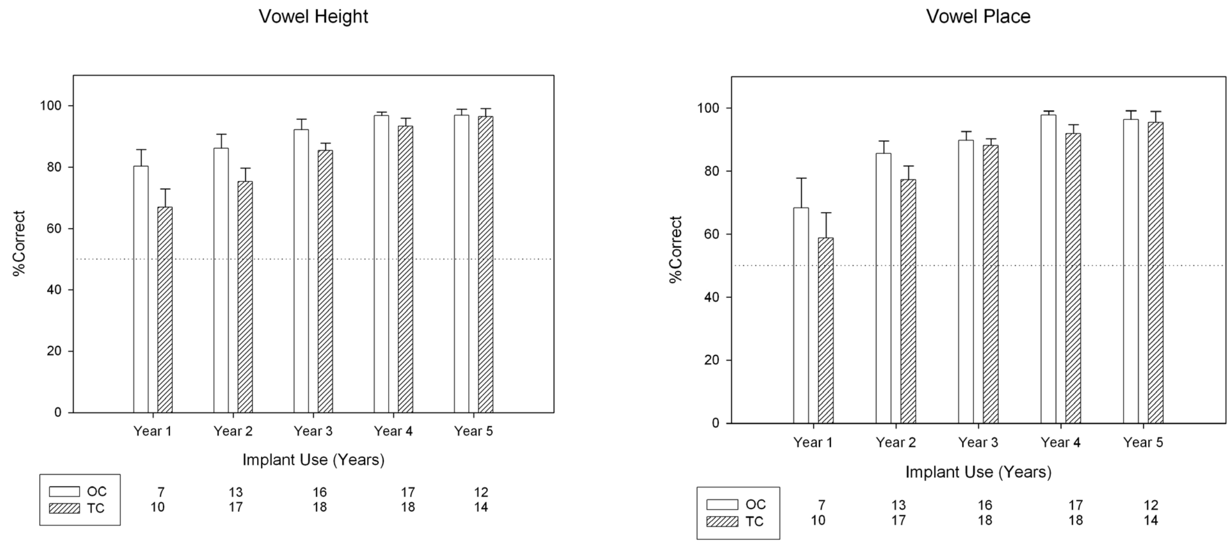


Figure 3. Percent correct for vowel features in the Minimal Pairs Test. Numbers below each graph represent sample size for the Oral Communication and Total Communication groups at each year of implant use. Error bars represent standard error.

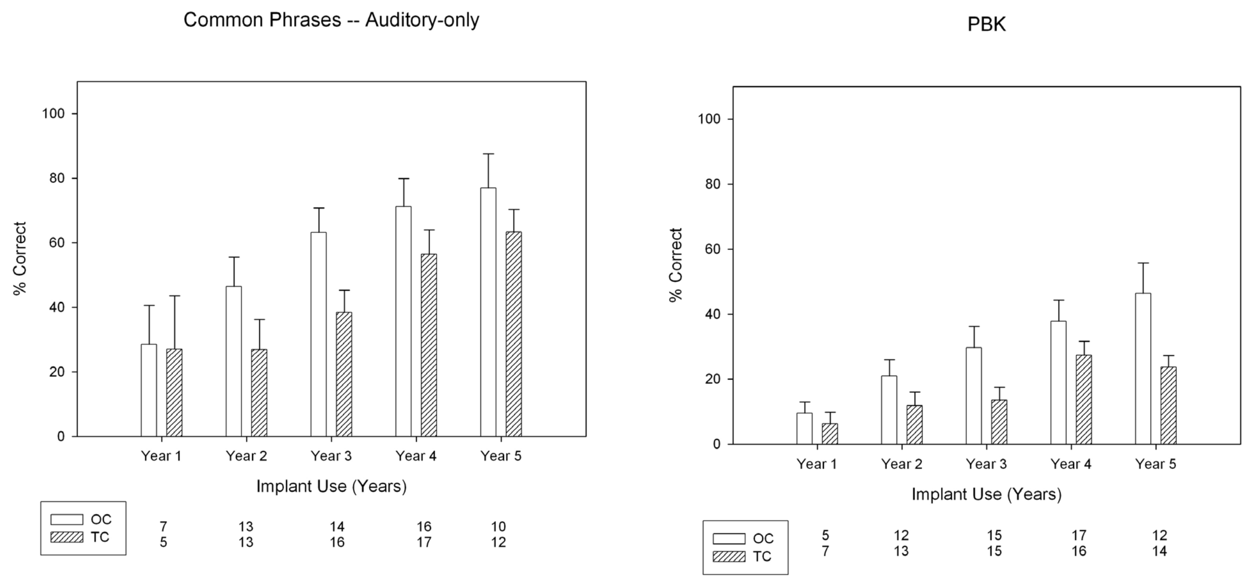


Figure 4. Percent correct for CPA and PBK. Numbers below each graph represent sample size for the Oral Communication and Total Communication groups at each year of implant use. Error bars represent standard error.

Minimal Pairs Test

	Oral Communication					**Total Communication**				
	Consonants			*Vowels*		*Consonants*			*Vowels*	
	Manner	Place	Voice	Height	Place	Manner	Place	Voice	Height	Place
	r =	r =	r =	r =	r =	r =	r =	r =	r=	r=
Speech Feature Discrimination										
Minimal Pairs – Place	.81**	--	--	--	--	.62**	--	--	--	--
Minimal Pairs – Voicing	.50*	.30	--	--	--	.44*	.36	--	--	--
Minimal Pairs – Vowel Height	.55*	.61**	.33	--	--	.34	.17	.36	--	--
Minimal Pairs – Vowel Place	.33	.57*	.31	.75**	--	.44*	.36	.43*	.15	--
Spoken Word Recognition										
PBK	.65**	.58*	.51*	.52*	.64**	.36	.43	.56*	.58*	.41
Sentence Comprehension										
Common Phrases – A	.50*	.58*	.46*	.47*	.71**	.38	.24	.63**	.45*	.41
Common Phrases – V	.30	.01	.33	.26	.08	-.11	-.43	-.04	-.19	.10
Common Phrases – A+V	.76**	.63**	.37	.47*	.43	.14	.10	.33	.47*	.12

* p < .05, ** p < .01 (n = 14-18)

Table 2. Correlations with the Minimal Pairs Test for Oral Communication and Total Communication groups at Year 3 of implant use.

General Discussion

Speech feature discrimination skills in deaf children with cochlear implants do improve across time, but may not reach the levels typically observed in normal-hearing children, particularly for minimal pairs involving discrimination of fine phonetic features such as consonant place and consonant voicing. By two years post-implant, the majority of these children achieved near-ceiling levels of discrimination performance for vowel height, vowel place, and consonant manner. Most of the children also achieved plateaus in performance but did not reach ceiling performance for consonant place and voicing. Communication mode affected outcome measures, with a slight advantage for Oral Communication children over Total Communication children in speech feature discrimination, spoken word recognition, and sentence comprehension. If speech feature discrimination underlies spoken word recognition and sentence comprehension, we would expect positive correlations between performance on the MP test and measures of word recognition and sentence comprehension (e.g., Blamey et al., 2001; Rabinowitz et al., 1992). As expected, we found that children who performed better on the MP test also performed better on more complex speech and language measures that rely on more basic speech feature discrimination skills. We also found that the relationship between speech feature discrimination, spoken word recognition, and sentence comprehension was stronger for children in oral-only communication environments compared with children in total communication environments. Moreover, these skills appear to increase in parallel rather than hierarchically (i.e., speech feature discrimination → spoken word recognition → sentence comprehension). Correlations among measures of speech feature discrimination, spoken word recognition, and sentence comprehension are consistent with the recent proposal that individual differences and variability in performance may arise not only from traditional demographic variables but also from central cognitive or psychological processes such as perception, learning, and memory (Pisoni et al., 1999).

References

Blamey, P.J., Sarant, J.Z., Paatsch, L.E., Barry, J.G., Bow, C.P., Wales, R.J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (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.

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

Pisoni, D.B., Cleary, M., Geers, A.E., & Tobey, E.A. (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *The Volta Review, 101,* 111-164.

Osberger, M.J., Miyamoto, R.T., Zimmerman-Phillips, S., Kemink, J.L., Stroer, B.S., Firszt, J.B., & Novak, M.A. (1991). Independent evaluation of the speech perception abilities of children with the Nucleus-22 channel cochlear implant system. *Ear & Hearing, 12,* S66-S80.

Rabinowitz, W.M., Eddington, D.K., Delhorne, L.A., & Cuneo, P.A. (1992). Relations among different measures of speech reception in subjects using a cochlear implant. *Journal of the Acoustical Society of America, 92,* 1869-1881.

Robbins, A.M., Renshaw, J.J., Miyamoto, R.T., Osberger, M.J., & Pope, M.J. (1988). *Minimal Pairs Test*. Indianapolis, IN: Indiana University School of Medicine.

Effects of Response Format in Spoken Word Recognition Tests: Speech Intelligibility Scores Obtained from Open-Set, Closed-Set, and Delayed Response Tasks[1]

Cynthia G. Clopper and David B. Pisoni

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

[1] This work was supported by NIH-NIDCD Research Grant DC00111 and Training Grant DC00012 to Indiana University.

Effects of Response Format in Spoken Word Recognition Tests: Speech Intelligibility Scores Obtained from Open-Set, Closed-Set, and Delayed Response Tasks

Abstract. In studies of spoken word recognition, factors such as the materials to be recognized and the degradation of the speech signal have been discussed at length in the literature. The factor that has been largely ignored, however, is that of response format. In particular, word recognition studies in the last forty years have been largely closed-set tasks, despite the fact that the original tests used by military personnel for testing radio equipment during World War II were designed to be open-set. Recent findings suggest that open-set and closed-set tests may differ in more than just their level of chance performance. Sommers, Kirk and Pisoni (1997) found that lexical competition and talker variability do not produce reliable effects on word recognition scores in closed-set tasks, but they do produce robust effects on performance in open-set tasks. The present study was designed to replicate the earlier findings reported by Sommers et al. and explore the response demands in a modified closed-set task. Specifically, one group of listeners participated in a closed-set word recognition task in which the response alternatives were not provided until one second after the offset of the signal. Their performance was compared to listeners in an open-set task and a traditional closed-set task. Results revealed effects of lexical competition only for the listeners in the open-set task. We assume that in open-set tests, the listener must access the lexicon. These findings therefore suggest that even a delay of one second is not adequate to force the listener to process the stimuli in a way that is similar to open-set processing. The major implication of these results is that even modified closed-set tasks cannot replace open-set tasks as valid measures of word recognition performance.

Introduction

Word recognition tasks have been used for more than a half-century in settings as diverse as the testing of military radio equipment (Miller, 1946a), studies of speech intelligibility (Horii, House, & Hughes, 1970), clinical tests of the auditory capabilities of hearing-impaired individuals (Owens, Kessler, Telleen, & Schubert, 1981), and studies of the cognitive implications of word superiority effects in letter recognition tasks (Reicher, 1969; Wheeler, 1970).

In the word recognition literature, factors such as mode of recognition (visual or auditory), the materials to be recognized, the content of the foils presented as response alternatives, and the type of masking or degradation have been manipulated by researchers to better suit the particular goals of their experiments. In spoken word recognition, Miller (1946a) preferred to use nonsense syllables because he was most interested in the accurate transmission of segmental information in speech. However, he also found that when using a limited set of real English words, speech intelligibility performance increased if the words were multisyllabic and phonetically distinct (Miller, 1946b).

House, Williams, Hecker, and Kryter (1965) used sets of phonetically similar real words because they were interested in how well listeners could differentiate words in English based on single consonant differences. Foster and Haggard (1987) also used real words but selected response alternatives based on binary feature distinctions instead of mere consonantal contrasts. More recently, Pisoni and his colleagues (Kirk, Pisoni, & Osberger, 1995; Luce & Pisoni, 1998; Meyer & Pisoni, 1999; Mullennix, Pisoni, &

Martin, 1989) have shown that the lexical properties of the target items (e.g., frequency, familiarity, and neighborhood density) as well as talker variability can affect spoken word recognition performance in both normal-hearing and hearing-impaired listeners.

Miller (1946a, b) could rely on the noise introduced through the radio equipment to degrade his stimuli, but noisy conditions have been simulated for normal-hearing listeners in other studies using a wide variety of methods including white noise (Sommers et al., 1997), envelope-shaped white noise (Horii et al., 1970), simulated airplane noise (Black, 1957), multi-talker babble (Nakanishi, 1988), and bit-flipping (Saldana, Pisoni, Fellowes, & Remez, 1996).

While these signal manipulations, as well as the clinical differences between normal-hearing and hearing-impaired populations, have been studied and discussed extensively in the literature, the issue of the nature of the task itself has received relatively little attention. Most of the early speech intelligibility tests described by Miller (1946a) used open-set tests of word (or syllable) recognition. However, other researchers began to switch to multiple choice speech intelligibility tests for speed and efficiency (Black, 1957). The underlying assumption of the new closed-set tests was that the process of word recognition would be the same, regardless of the response format of the test. It was thought that the only difference between open-set and closed-set tasks was chance performance ($1/\infty$ in open-set tests; 1/N in closed-set tests, where N is the number of response alternatives).

In closed-set tests, researchers have varied the response alternatives in consideration of phonological properties to examine more closely the effects of linguistic properties such as phoneme confusability on word recognition. In designing one of the earliest closed-set speech intelligibility tests, Black (1957) selected three foils for each target word from incorrect responses to the targets presented in an open-set condition in noise. That is, he used foils that he knew to be confusable with his targets based on data collected in a similar task with human participants. For example, a target word *burst* might be accompanied by the foils *nurse*, *first*, and *birth* on the response sheet.

House et al. (1965) used a different approach to create their six-alternative forced choice speech intelligibility test. They created 50 sets of six CVC words each. In 25 of the sets, all of the words in each set contained the same initial consonant and the same vowel. For example, *bat*, *bad*, *back*, *bass*, *ban*, and *bath* are one set in which only the final consonant varies between the words. In the remaining 25 sets, all of the words contained the same vowel and the same final consonant. An example of this kind of set might include *led*, *shed*, *red*, *bed*, *fed*, and *wed* in which all of the words rhyme but have different initial consonants. All of the foils then differed from the target word by only a single consonant (either the initial or the final). The result of this kind of design is that any word in a given set could be used as the target word and the other words in the set would be its foils.

In the Minimal Auditory Capabilities (MAC) battery, Owens et al. (1981) used a similar design for determining phoneme discrimination abilities. In a series of four-alternative forced choice tests, Owens et al. investigated word recognition abilities for English CVC words differing either in initial consonant (e.g., *din*, *bin*, *fin*, and *gin*), final consonant (e.g., *rid*, *rip*, *rib*, and *ridge*), or vowel (e.g., *fool*, *full*, *fall*, and *foul*).

Finally, Foster and Haggard (1987) considered an even smaller unit of linguistic contrast in their creation of the Four Alternative Auditory Feature (FAAF) test. Instead of varying their foils and targets based on phonemic consonantal contrasts, Foster and Haggard built lists of targets and foils that differed only on individual featural dimensions based on minimal pairs. For example, the target word *mail* (with an initial bilabial nasal) might be accompanied by the foils *bail* (with an initial bilabial stop), *nail* (with an initial alveolar nasal), and *dale* (with an initial alveolar stop).

Despite the many systematic considerations of the content of the targets and the foils, the effects of response format on spoken word recognition performance have not been considered. Pollack, Rubenstein, and Decker (1959) reported that differences in word frequency effects were observed between known (closed-set) and unknown (open-set) word sets in speech intelligibility tests. In particular, word frequency affected performance on the unknown sets of words but not the known sets of words. This effect was found regardless of the number of words in the set, which ranged in size from 8 to 144 words. These early results suggest that some aspects of the normal word recognition process, such as those responsible for word frequency effects in recognition, may be bypassed in closed-set speech intelligibility tests.

In a recent study from our laboratory, Sommers et al. (1997) used both an open-set and a closed-set task in evaluating the word recognition performance of cochlear implant users. They manipulated lexical competition (based on neighborhood density and neighborhood frequency), talker variability, and response format in both normal-hearing listeners (in both quiet and noisy conditions) and cochlear implant users (in a quiet condition only). Sommers et al. found that performance on the closed-set task was better than performance on the open-set task for both groups of listeners. However, the effects of lexical competition and talker variability were observed only in the open-set condition. These effects were found in both groups of listeners. No significant effects were found for the normal-hearing listeners in the quiet condition because their performance was at ceiling for all variables.

The findings by Sommers et al. (1997) also suggest that the processes used in recognizing spoken words in closed-set test formats may not be equivalent to the processes used in open-set word recognition. In particular, it is assumed in most models of word recognition that the lexicon is accessed through some kind of activation process based on the properties of the acoustic signal (Jusczyk & Luce, 2002). While open-set tasks using isolated single words may not be a perfect model of real-life language situations that typically involve the use of sentences, the contribution of semantic context, and lower levels of degradation, closed-set word recognition tasks may fundamentally bypass some of the higher-level lexical access processes that are certainly at work in everyday language situations. Under these listening conditions factors such as lexical competition and talker variability may come into play.

Given that closed-set tests are easier to administer and score, it would be useful in clinical settings to have a closed-set test to measure word recognition performance. However, given that open-set tests may employ a fundamentally different set of cognitive processes, scores on closed-set tests in clinical situations might provide an inflated measure of the listener's speech perception skills. The goal of the present study was to determine if a modified closed-set task could produce the same effects of lexical competition and talker variability as an open-set test. In particular, the closed-set task used in this study was modified so that the listeners did not see the response alternatives until 1000 ms after the presentation of the spoken test word. We predicted that this delay in the response alternatives might result in the activation of the normal (i.e., open-set) word recognition process, thus revealing the same effects found in open-set tests. One group of young normal-hearing listeners participated in the delayed closed-set condition. Two additional groups of listeners were used for comparison: one was assigned to an open-set condition and the other was assigned to a more traditional closed-set condition in which the response alternatives were provided before the presentation of the stimulus word.

Experiment 1: Level of Signal Degradation

In order to determine appropriate levels of signal degradation, a pilot study was carried out using three levels of signal degradation in two conditions: open-set and closed-set. The goal of this preliminary study was simply to determine which levels of signal degradation would allow for performance above floor in the open-set condition and below ceiling in the closed-set condition.

Methods

Materials. A set of 132 CVC English words were selected for use in this study from the Modified Rhyme Test (MRT; House et al., 1965) and the Phonetically Balanced (PB; IEEE, 1969) word lists. The words were divided into two groups based on measures of lexical competition ('easy' and 'hard') with 66 words in each group. Based on the word counts in Kučera and Francis (1967), mean log frequency was equated across the two groups of words. Similarly, mean word familiarity, as judged by Indiana University undergraduates (Nusbaum, Pisoni, & Davis, 1984), was also equated across the two groups. The lexically 'hard' words had a significantly higher mean lexical density (defined as all words that differ from the target by a single phoneme substitution, deletion, or addition) and the neighbors of the 'hard' words had a significantly higher mean log frequency than the neighbors of the lexically 'easy' words. As in Sommers et al. (1997), the 'easy' and 'hard' sets of test words were defined based on these differences in mean density and mean neighborhood log frequency. Table 1 shows a summary of the means for each set of words on the four lexical measures, as well as significance values for the t-tests run to compare the means.

	Easy	**Hard**	
Mean Log Frequency	1.91	2.02	$p = .41$
Mean Familiarity	6.81	6.70	$p = .27$
Mean Density	16.03	24.86	$p < .001$
Mean Neighborhood Log Frequency	1.83	2.17	$p < .001$

Table 1. Lexical properties of 'easy' and 'hard' words.

Five male talkers were selected from a total of 20 talkers (10 males and 10 females) who were recorded reading the MRT and PB word lists for the PB/MRT Word Multi-Talker Speech Database (Speech Research Laboratory, Indiana University). Based on the results of a similarity judgment task (Goh, 2001), a clustering analysis was computed for the male and female talkers. In order to ensure that the selected talkers could be discriminated from one another, two talkers were taken from two of the major clusters and the fifth was taken from the third cluster in the clustering solution. In the two-dimensional multi-dimensional scaling solution to the similarity judgments, the five selected talkers were well-dispersed in both dimensions of the space.

Each of the 132 words was spoken by each of the five talkers, for a total of 660 tokens. Each token was stored in an individual sound file in .wav format. For the present study, the tokens were degraded using a bit-flipping procedure written in Mathworks Matlab. In bit-flipping, degradation is introduced into the signal by flipping the sign of a randomly selected proportion of the bits in the signal. The higher the percentage of bits that are flipped, the more degraded the signal is. In this case, three levels of degradation were selected: 10%, 20%, and 30%.

For the closed-set condition, a six-alternative forced-choice task was designed. Each of the five foils for each target word differed from the target by the substitution of a single phoneme. Ideally, all of the foils were rated by undergraduates as having a familiarity rating of greater than 6.0 on a 7-point scale (Nusbaum et al., 1984). In addition, the goal was to have two foils higher in frequency, two foils lower in frequency, and one foil the same frequency as the target. Finally, two foils differed from the target with respect to the initial consonant, two with respect to the final consonant, and one with respect to the vowel. This general schema for selecting foils could not be followed in all cases due to the constraints of the English language. Therefore, some sets of foils did not match the schema with respect to minimum familiarity, frequency distributions, or phoneme substitution location, but in all cases the foils differed from the targets by the substitution of a single phoneme and the familiarity of all foils was greater than 5.0 based on scores obtained from the Hoosier Mental Lexicon database (Nusbaum et al., 1984). A complete list of targets and foils is shown in Appendix A.

Listeners. Twenty-one Indiana University undergraduates, 10 males and 11 females, participated as listeners in this experiment. Ten listeners were assigned to the open-set condition and 11 listeners were assigned to the closed-set condition. All listeners were native speakers of English, with no history of hearing or speech disorders reported at the time of testing. The listeners received partial course credit for participating in the experiment.

Procedure. The listeners were seated at personal computers equipped with Beyerdynamic DT100 headphones. All 132 words were presented one time at each of the three levels of degradation, for a total of 396 trials per listener. All of the words at one level of degradation were presented in random order within a single block. Presentation order of the degradation blocks was balanced across listeners. Each listener heard only one of the five talkers and that talker remained constant across all three blocks. The talkers were randomly assigned to the listeners such that there were two or three listeners per talker per listener group.

The listeners assigned to the open-set condition heard the words presented one at a time over the headphones at 75dB and were simply asked to type in the word that they thought they had heard using a standard keyboard. The next trial was initiated by pressing a "Next Trial" button on the computer screen by clicking on it with the mouse.

The listeners assigned to the closed-set condition also heard the words presented one at a time over the headphones at 75dB. Simultaneously with the onset of the auditory presentation of each word, six response alternatives were presented on the computer screen. After hearing the word, listeners were asked to use the mouse to select which one of the six response alternatives they thought they had heard. The next trial was initiated by pressing a "Next Trial" button on the computer screen by clicking on it with the mouse.

Prior to data analysis, open-set responses were corrected by hand for obvious typographical errors in cases where the response given was not a real word and for homonyms such as *pear* for *pair*.

Results

Figure 1 shows the mean performance by the two groups of listeners for each of the three levels of degradation. A two-way ANOVA (degradation x response format) confirmed the significant effect of degradation ($F(2, 60) = 68.4$, $p < .001$). Post-hoc Tukey tests revealed that performance at 10% degradation was significantly better than performance at 20% and 30% degradation ($p < .001$ for both) and that performance at 20% degradation was significantly better than performance at 30% degradation ($p < .001$). In addition, there was a significant main effect of response format ($F(1, 60) = 350.3$, $p < .001$). Closed-set performance was always better than open-set performance, as expected based on the results of Sommers et al. (1997). The interaction between response format and level of degradation was not significant.

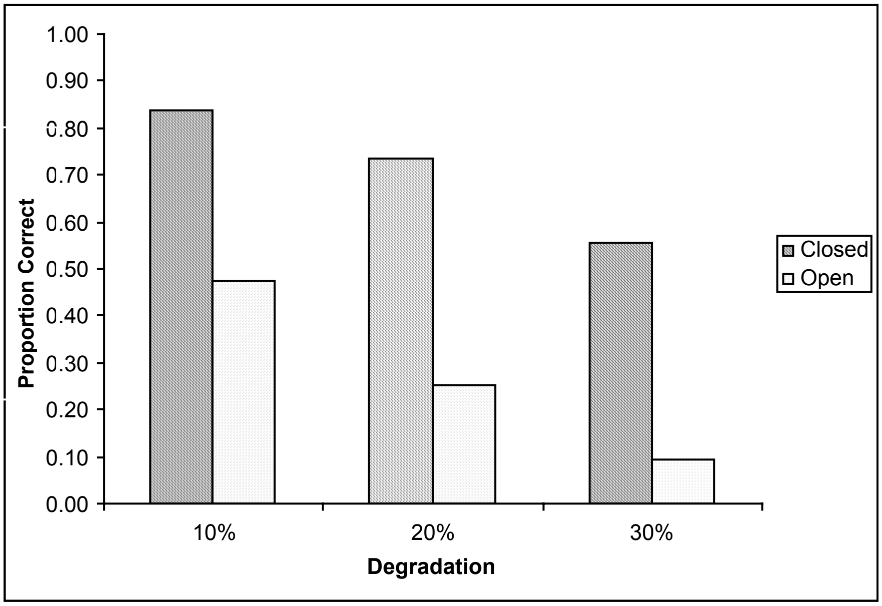


Figure 1. Word recognition performance as a function of response format and level of degradation.

Discussion

Given that open-set performance was close to floor (9%) at 30% degradation and that closed-set performance was not at ceiling (84%) at 10% degradation, 10% and 20% levels of degradation were selected for use in the next experiment which manipulated the closed-set response format.

Experiment 2: Response Format Task Analysis

The goal of this study was to determine if a modified closed-set test of speech intelligibility could be used as an alternative to open-set tests in clinical settings. Therefore, a delayed response closed-set task ('after' condition) was used as the test condition. In addition, open-set ('open' condition) and traditional closed-set ('before' condition) tasks were used as control conditions. Response condition was a between-subject variable. Other variables that have been shown to reveal differences between open-set and closed-set word recognition tests were manipulated within subjects: level of signal degradation, talker variability, and lexical competition.

Methods

Materials. The same test words, spoken by the same five male talkers in Experiment 1, were used in this experiment. In addition, the same response alternatives were used in the two closed-set conditions as in Experiment 1. In this experiment, only 10% and 20% levels of degradation were used.

Listeners. Eighty-one Indiana University undergraduates, 26 males and 53 females, served as listeners in this experiment. All received partial course credit for participating. Data from nine listeners were excluded from the final data analysis because the listeners were bilingual (four participants), were non-native speakers of English (one participant), had a history of speech or hearing disorder (three participants), or fell outside the age distribution of the other participants (one participant). Data from 24 listeners in each of the three conditions are reported in the final analyses below. All 72 of these listeners were monolingual native speakers of English with no reported history of hearing or speech disorders.

Procedure. The listeners were seated at personal computers equipped with Beyerdynamic DT100 headphones. Each of the 132 words were presented one time to each listener. The procedure was divided into a practice block and four test blocks. The practice block consisted of two randomly selected 'easy' words and two randomly selected 'hard' words spoken by the same talker presented at 10% degradation. Listeners completed the four practice trials and were encouraged to ask questions regarding the procedure before continuing on to the test blocks. Each of the four test blocks contained 16 'easy' and 16 'hard' words. The four blocks represented all possible combinations of level of degradation (10% or 20%) and number of talkers (single or multiple). For each listener, one talker was selected as the single talker and that talker was only used in the practice block and the two single talker blocks. The remaining four talkers were used in the multiple talker blocks. The words in each block were randomly and exhaustively selected for each listener, such that each listener heard each word only once, but the talker and block in which each word appeared was random. Single talkers and block orders were balanced across listeners to ensure that any observed effects were not due to characteristics of a specific talker.

As in Experiment 1, the listeners in the open-set condition heard the words presented one at a time over headphones at 75dB and were asked to type in the word that they thought they had heard using a standard keyboard. The next trial was initiated by pressing a "Next Trial" button on the computer screen by clicking on it with the mouse.

The listeners in the 'before' closed-set condition heard the words presented one at a time over headphones at 75dB. One second prior to the onset of the auditory presentation of the words, the six response alternatives were presented in random order in a single row on the screen. After hearing the word, listeners were asked to use the mouse to select which one of the six response alternatives they thought they had heard. The next trial was initiated by pressing a "Next Trial" button on the computer screen by clicking on it with the mouse.

The listeners in the 'after' closed-set condition also heard the words presented one at a time over headphones at 75dB. However, in this condition, the six response alternatives were presented on the screen one second after the end of the auditory presentation of the words. After the presentation of the response alternatives, listeners were asked to use the mouse to select which one of the six response alternatives they thought they had heard. The next trial was initiated by pressing a "Next Trial" button on the computer screen by clicking on it with the mouse.

Prior to the analysis of the data, open-set responses were corrected by hand for obvious typographical errors and homonyms, using the same criteria as in Experiment 1.

Results

Response Format. We observed an overall effect of response format on performance, collapsed across all of the other independent variables: level of degradation, talker variability, and lexical competition, as shown in Figure 2. A one-way ANOVA revealed a significant main effect of response format ($F(2, 285) = 279.6$, $p < .001$). Post-hoc Tukey tests revealed that performance in the open-set condition was significantly worse than performance in either of the closed-set conditions ($p < .001$ for both) and that performance in the 'after' condition was significantly worse than in the 'before' condition ($p < .001$).

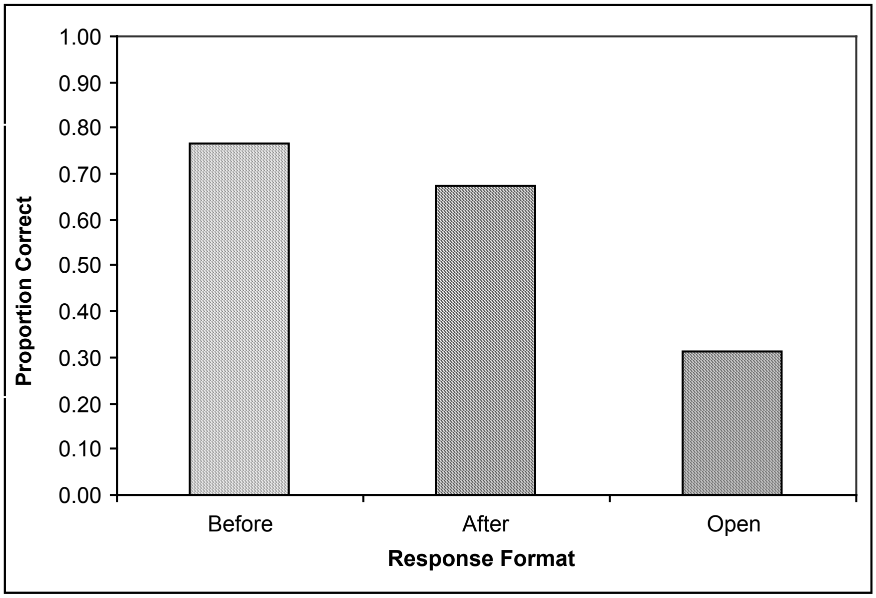


Figure 2. Word recognition performance as a function of response format, collapsed across level of degradation, talker variability, and lexical competition.

Signal Degradation. As shown in Figure 3, an effect of level of degradation was observed for all three response formats, collapsed across talker variability and lexical competition. A two-way ANOVA (response format x degradation) revealed a significant main effect of response format ($F(2, 141) = 479.0$, $p < .001$). Post-hoc Tukey tests again confirmed that performance in the 'before' condition was significantly better than performance in the 'after' and open-set conditions ($p < .001$ for both) and that performance in the 'after' condition was better than in the open-set condition ($p < .001$). There was also a significant main effect of degradation ($F(1, 141) = 235.3$, $p < .001$). Performance at low degradation (10%) was better than performance at high degradation (20%). Planned post-hoc t-tests revealed a significant effect of degradation for all three conditions: $t(23) = 9.5$, $p < .001$ for the 'before' condition; $t(23) = 10.4$, $p < .001$ for the 'after' condition; $t(23) = 15.5$, $p < .001$ for the open-set condition. The response format x signal degradation interaction was also significant ($F(2, 1) = 4.7$, $p = .01$). In order to determine the location of the interaction, a one-way ANOVA on the mean differences in performance between the two levels of degradation for the three conditions was calculated. The ANOVA revealed a significant effect of condition ($F(2, 69) = 8.2$, $p < .001$). Post-hoc Tukey tests revealed that the effect of degradation was significantly lower for the 'before' condition than either the open-set condition or the 'after' condition ($p < .001$ and $p < .05$, respectively). The effect was not significantly different between the 'after' condition and the open-set condition.

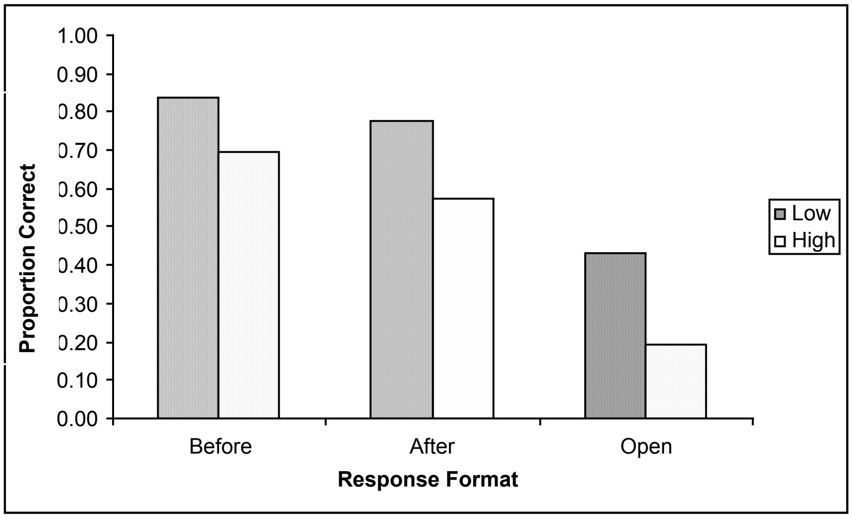


Figure 3. Word recognition performance as a function of response format and level of degradation, collapsed across talker variability and lexical competition.

Talker Variability. Unexpectedly, the effect of talker variability was not significant for any of the three conditions, as shown in Figure 4. A two-way ANOVA (response format x talker variability) was performed, collapsed across level of degradation and lexical competition. The main effect of response format was again significant ($F(2, 141) = 413.2$, $p < .001$). The main effect of talker variability, however, was only marginally significant ($F(1, 141) = 2.8$, $p = .09$), although the difference between single and multiple talker blocks was in the predicted direction. Planned post-hoc t-tests revealed a non-significant difference for all three conditions: $t(23) = .5$, $p = .62$ for the 'before' condition; $t(23) = 1.48$, $p = .15$ for the 'after' condition; $t(23) = 1.31$, $p = .20$ for the open-set condition. The response format x talker variability interaction was not significant.

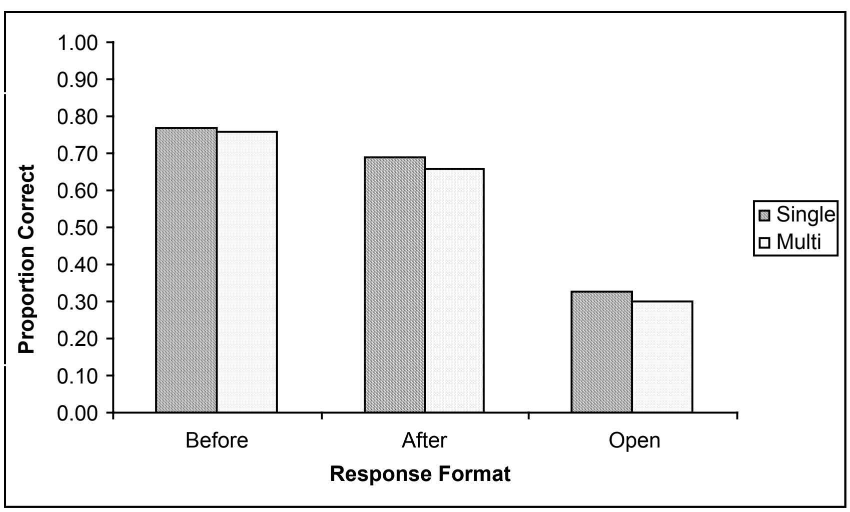


Figure 4. Word recognition performance as a function of response format and talker variability, collapsed across level of degradation and lexical competition.

Lexical Competition. The effect of lexical competition was significant only for the open-set condition, as shown in Figure 5. A two-way ANOVA (response format x lexical competition), collapsed across level of degradation and talker variability, again confirmed a significant main effect of response format ($F(2, 141) = 623.0$, $p < .001$). The main effect of lexical competition was not significant ($F(1, 141) = 3.4$, $p = .70$), but there was a significant response format x lexical competition interaction ($F(2, 1) = 4.3$, $p = .02$). Planned post-hoc t-tests revealed a significant effect of lexical competition only in the open-set condition: $t(23) = -.96$, $p = .35$ for the 'before' condition; $t(23) = .73$, $p = .48$ for the 'after' condition; $t(23) = 4.67$, $p < .001$ for the open-set condition.

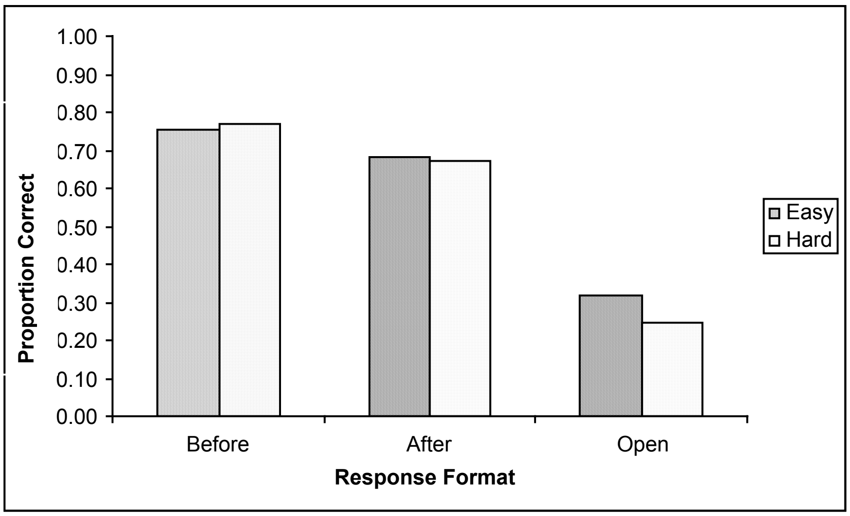


Figure 5. Word recognition performance as a function of response format and lexical competition, collapsed across level of degradation and talker variability.

Discussion

As expected, word recognition performance was affected by the signal degradation in all conditions. Unexpectedly, we did not find effects of talker variability in any of the conditions in the experiment. In previous studies, talker variability effects have been found in open-set tasks for normal-hearing listeners in noisy conditions (Sommers et al., 1997), hearing-impaired listeners in quiet conditions (Sommers et al., 1997), and normal-hearing listeners for lexically 'easy' and lexically 'hard' words (Mullennix et al., 1989). A more detailed analysis of the responses revealed that 12 of the 24 listeners in the 'before' condition, 14 of the 24 listeners in the 'after' condition, and 13 of the 24 listeners in the open-set condition performed better on the single talker blocks than the multiple talker blocks. The absence of an effect of talker variability is therefore not due to a consistent, but small, performance difference across listeners. Rather, it is due to the fact that nearly half of the participants showed an improvement in the opposite direction.

This difference between the current results and those found in the previous studies may be due to differences in experimental design. Unlike Sommers et al. (1997) who used five male and five female talkers or Mullennix et al. (1989) who used 15 different talkers, we used only four male talkers in the multi-talker blocks. It is possible that the absence of a talker effect could be due to either the small number of talkers used in the multi-talker blocks or the fact that the talkers were all male. In addition, the process of degradation using the bit-flipping algorithm may have resulted in making the talkers perceptually more similar than the degradation used in other studies, and therefore reduced the talker variability effect. Both Sommers et al. and Mullennix et al. used white noise to reduce performance. Finally, in this experiment talker-variability was treated as a within-subjects variable. In Mullennix et al. and Sommers et al., it was treated as a between-subjects variable. Obviously, further studies investigating these response conditions are needed to confirm under which conditions talker variability effects occur and to determine whether the absence of the effect of talker variability in the present study was due to an effect of the number or gender of the talkers, the type of degradation used, or the design of the experiment itself with respect to between- and within-subject variables.

As predicted, however, the effects of lexical competition were found only in the open-set condition, confirming that even when the response alternatives are presented to listeners after a delay of one second, lexical access still occurs through a process that may be fundamentally different than the one used in open-set situations. The present set of findings confirms the need for the use of open-set tests in clinical situations and suggests that even a delayed closed-set format may not be a suitable alternative to open-set tests of speech perception and spoken word recognition. It is possible that a longer delay before the response alternatives are presented or a more complex divided attention task might reveal the same effects as an open-set task.

Conclusions

Earlier studies have reported that lexical competition effects are not observed in closed-set word recognition tasks. The present investigation replicated and confirmed those results and provided further evidence for a disparity in performance between open-set and closed-set speech intelligibility tests. In particular, even a modified closed-set test involving a delay before the presentation of the response alternatives did not reveal lexical competition effects in normal-hearing listeners under noisy conditions. Therefore, this kind of delayed closed-set format may not be any more useful in clinical situations than the traditional closed-set tasks with simultaneous presentation of the stimulus and the response alternatives. This study does, however, contribute to the growing literature on the processes involved in word recognition and suggests that higher level processes, such as lexical access that are affected by the lexical properties of the test words can be deferred in favor of other processes such as pattern-matching, even over a delay of 1000 ms.

References

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

Foster, J.R. & Haggard, M.P. (1987). The Four Alternative Auditory Feature test (FAAF) – Linguistic and psychometric properties of the material with normative data in noise. *British Journal of Audiology, 21*, 165-174.

Goh, W. (2001). Talker variability effects in short-term and long-term memory for spoken words. Doctoral dissertation, Indiana University.

Horii, Y., House, A.S., & Hughes, G.W. (1970). A masking noise with speech-envelope characteristics for studying intelligibility. *Journal of the Acoustical Society of America, 49*, 1849-1856.

House, A.S., Williams, C.E., Hecker, M.H.L., & Kryter, K.D. (1965). Articulation-testing methods: Consonantal differentiation with a closed-set response. *Journal of the Acoustical Society of America, 37*, 158-166.

IEEE. (1969). IEEE recommended practice for speech quality measurements. (IEEE Report No. 297).

Jusczyk, P.W. & Luce, P.A. (2002). Speech perception and spoken word recognition: Past and present. *Ear & Hearing, 23*, 2-40.

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.

Kučera, F. & Francis, W. (1967). *Computational Analysis of Present-Day American English*. Providence, RI: Brown University Press.

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

Meyer, T.A. & Pisoni, D.B. (1999). Some computational analyses of the PBK test: Effects of frequency and lexical density on spoken word recognition. *Ear & Hearing, 20*, 363-371.

Miller, G.A. (1946a). Articulation testing methods. In *Transmission and Reception of Sounds under Combat Conditions.* Summary Technical Report of Division 7, NDRC. Washington, D.C. pp. 69-80.

Miller, G.A. (1946b). Intelligibility of speech: Special vocabularies. In *Transmission and Reception of Sounds under Combat Conditions.* Summary Technical Report of Division 7, NDRC. Washington, D.C. pp. 81-85.

Mullennix, J.W., Pisoni, D.B., & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85*, 365-378.

Nakanishi, Y. (1988). Competing noise for the evaluation of hearing-aid performance: Comparison of multi-talker babble vs. speech-shaped noise. *RIEEC Report*, *37*, 9-20.

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

Owens, E., Kessler, D.K., Telleen, C.C., & Schubert, E.D. (1981). The Minimal Auditory Capabilities (MAC) Battery. *Hearing Aid Journal*, *34,* 9-34.

Pollack, I., Rubenstein, H., & Decker, L. (1959). Intelligibility of known and unknown message sets. *Journal of the Acoustical Society of America, 31*, 273-279.

Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 275-280.

Saldana, H.M., Pisoni, D.B., Fellowes, J.M., & Remez, R.E. (1996). Audio-visual speech perception without speech cues. *Proceedings of the International Conference on Speech and Language Processing, 96,* 2187-2190.

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

Wheeler, D.D. (1970). Processes in word recognition. *Cognitive Psychology, 1*, 59-85.

Appendix A

Easy Words

Target	Foils				
beach	teach	reach	batch	beef	beat
book	shook	took	back	bull	bush
bud	thud	dud	bid	bum	buck
buff	puff	cuff	beef	bug	bum
cave	pave	shave	curve	cage	cane
check	deck	neck	choke	chuck	chair
coil	soil	boil	curl	coal	coin
cub	pub	tub	cob	cuff	cup
cud	mud	dud	could	cuff	cub
cuff	muff	puff	calf	cub	cut
dab	gab	jab	dub	dash	dad
death	den	deaf	debt	deck	dead
deck	neck	check	dock	debt	den
dike	bike	tyke	duke	dine	dial
dim	rim	him	dome	dish	did
dub	rub	pub	dab	dud	does
dug	thug	chug	dog	dutch	does
fang	pang	tang	gang	fad	fat
fern	churn	yearn	faun	fan	firm
fib	rib	bib	fizz	fish	fit
fig	jig	big	fog	fill	fit
gang	hang	rang	gong	gag	gas
heap	sheep	leap	hop	heal	heat
hire	tire	fire	hair	hike	hive
hive	chive	dive	heave	hike	hire
hook	cook	shook	hike	hoof	hood
hop	chop	mop	hoop	hip	hot
jam	lamb	dam	germ	jab	jack
job	rob	bob	jab	jot	jar
kid	hid	did	cod	kill	king
king	ring	thing	kick	kid	kill
lush	gush	mush	lash	luck	love
map	gap	nap	mop	mad	man
moose	deuce	noose	mace	mood	moon
mop	hop	chop	mope	mock	mob
nab	gab	jab	knob	nag	knack
neck	check	wreck	nick	knock	net
pass	gas	mass	peace	pal	pat
path	math	bath	patch	pat	pass
peach	reach	leach	perch	peal	peak
peas	seize	these	poise	peal	peat
perk	jerk	work	puck	pearl	perch
pig	rig	jig	peg	pin	pitch
pub	dub	cub	rub	pun	pup

puff	tough	cuff	pub	puck	putt
pup	cup	pip	peep	puck	pub
pus	bus	fuss	pass	pub	pun
rang	gang	sang	ring	ran	wrap
rib	fib	bib	rob	ring	rick
rise	size	guise	rose	rhyme	right
rouse	rout	raise	rose	rise	dowse
rub	tub	pub	robe	rum	rush
sag	lag	gag	sash	sack	sad
sang	hang	rang	sung	sack	sat
save	cave	gave	serve	sail	safe
shop	hop	top	ship	shock	shot
sob	cob	lob	sub	sod	psalm
soil	boil	royal	sail	sole	cell
sub	pub	rub	sob	suck	some
tab	gab	nab	tube	tang	tag
tang	pang	sang	tongue	tack	tab
team	theme	seam	term	tease	teeth
tease	cheese	peas	these	teach	team
top	hop	shop	tape	type	tar
turf	serf	tough	term	tern	Turk
wedge	hedge	ledge	wage	well	web

Hard Words

Target	Foils				
bait	date	gate	bite	beige	bathe
ban	man	fan	bean	bash	bang
bead	weed	knead	bed	beak	bean
beak	leak	reek	balk	bean	bead
bean	scene	lean	ban	beak	bead
bed	dead	red	bad	bell	bet
bill	pill	will	ball	bit	big
bought	sought	taught	bite	ball	boss
bun	pun	shun	burn	bug	bud
cake	fake	wake	coke	cane	came
den	ken	hen	dine	deaf	dare
dill	mill	hill	doll	dish	dip
dud	mud	thud	dude	dug	dub
fate	date	mate	foot	fail	fame
feel	kneel	deal	fail	full	feat
fill	dill	hill	fail	fizz	fig
fun	sun	gun	phone	fuss	fudge
gill	fill	sill	gull	ghoul	give
heal	feel	meal	hail	heed	heath
hid	lid	mid	hood	hip	hick
hip	chip	zip	heap	hill	hick
hot	lot	got	hit	hut	hop
kin	fin	shin	keen	kick	kit

kit	mitt	bit	cot	kin	kill
lace	mace	chase	lease	lame	lake
lame	maim	tame	limb	lane	lake
led	wed	red	load	let	less
mad	bad	pad	mood	mash	man
male	pail	hail	mull	main	make
mat	sat	rat	moat	math	mad
meat	seat	feat	might	meal	mean
muck	tuck	shuck	make	mull	mutt
neat	heat	peat	gnat	knead	niece
nut	shut	rut	gnat	null	none
pace	face	vase	pass	pail	pave
pale	wail	tale	pool	pain	page
pan	fan	tan	pin	patch	pat
peat	feat	seat	pout	peas	peep
pot	shot	lot	peat	palm	pop
puck	shuck	chuck	perk	putt	pub
pun	bun	done	pawn	puff	puck
race	base	face	rice	rain	rate
rake	bake	take	wreck	raid	rave
rat	gnat	hat	rut	rap	ran
rave	pave	shave	cave	rage	rate
ride	hide	tide	road	writhe	right
sack	hack	knack	soak	sad	sat
sad	mad	bad	side	sang	sap
sale	mail	vail	soil	save	sage
sane	cane	lane	sin	sake	sage
sap	yap	gap	sip	sash	sad
seed	weed	deed	side	seize	seat
seek	weak	beak	soak	seem	seat
seep	peep	reap	sit	siege	seat
sit	fit	hit	set	sill	sing
sun	done	one	sign	sub	such
tale	pail	mail	pool	tape	tame
tan	can	man	tune	tag	tang
teak	leak	beak	tyke	teach	teeth
teal	veal	real	tall	tease	teach
tier	leer	veer	tar	tease	teach
tot	jot	cot	tote	tight	tar
wed	fed	dead	wade	wedge	web
wick	hick	pick	work	witch	wig
wig	pig	dig	wag	witch	wick
wit	sit	mitt	watt	wing	with

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

Perception of Dialect Variation: Some Implications for Current Research and Theory in Speech Perception[1]

Cynthia G. Clopper and David B. Pisoni

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

[1] This work was supported by the NIH NIDCD R01 Research Grant DC00111 and the NIH NIDCD T32 Training Grant DC 00012 to Indiana University. We would also like to acknowledge the valuable advice that we received on various aspects of this project from Kenneth deJong, Caitlin Dillon, Luis Hernandez, and Robert Nosofsky, as well as the assistance of Jeffrey Reynolds and Adam Tierney with data collection.

Perception of Dialect Variation: Some Implications for Current Research and Theory in Speech Perception

Abstract. Despite the mounting evidence that variation and variability play an important role in spoken language processing, few speech researchers have investigated the relationship between dialect variation and human speech perception. Sociolinguists, on the other hand, have extensively documented linguistic variation and its social implications, but have largely ignored how dialect variation is perceived and encoded by naïve listeners. We review and discuss several different methodologies that have been used to study the perception of dialect variation. Data collected from map drawing tasks in sociolinguistics, matched-guise studies in social psychology, caricatures in forensic linguistics, and perceptual categorization in cognitive psychology have all contributed to our understanding of how linguistic variation is perceived, processed, encoded, and used by naive listeners in normal language situations. The implications for these findings for models of speech perception, speech recognition and speech synthesis technologies, and theoretical linguistics are discussed.

Introduction

Variability in speech comes in many forms: within-speaker variability, cross-speaker variability, segment realization variability, and word environment variability (Klatt, 1989). One approach to the study of speech perception and spoken language processing is to ignore these sources of variability and to work to normalize the signal to find acoustic invariances across utterances, talkers, and contexts. A different approach, however, is to recognize these sources of variability as a natural consequence of language variation and work to understand how variation and variability are processed and encoded in speech perception. This second alternative espouses the notion that variation matters and that listeners can and do encode the indexical properties of the talker as part of the normal speech perception process (Pisoni, 1993).

Fifty years ago, Peterson and Barney (1952) recorded 33 men, 28 women, and 15 children reading two lists of ten [hVd] syllables. They took first and second formant frequency measurements for each of the vowels produced by each of the talkers. A scatterplot of the F1 values by the F2 values for each talker revealed a vowel space containing large overlapping ellipses for each of the ten vowels. In their discussion, Peterson and Barney pointed out the continuous nature of the vowel space; there are not obvious breaks in the data as one moves from one phoneme to another in the F1 x F2 plane. In addition, they noted that the distribution of tokens for a single phoneme represents the enormous variability with which any given vowel is produced across different talkers.

More recently, Hillenbrand, Getty, Clark, and Wheeler (1995) and Hagiwara (1997) have replicated Peterson and Barney's (1952) findings with respect to individual talker variation in terms of [hVd] formant frequency measures. Both of these more recent studies also found differences in mean formant values across their talkers compared to the formant values in the Peterson and Barney study. In particular, Hillenbrand et al. found a dramatic shift in the low vowels of their talkers, reflecting the Northern Cities vowel shift that has taken place in the last 50 years in urban areas in the northern United States. Hagiwara, on the other hand, found a dramatic shift in the back vowels, reflecting the southern California trend of back vowel fronting. These two sets of results suggest that researchers who are interested in the study of human speech perception should consider not only the effects of talker variability on vowel formants in production, but also the impact of regional dialect variation on vowel production and the implications of these differences in spoken language processing.

While this acoustic-phonetic research in the speech sciences has been carried out, sociolinguists have conducted extensive research on vowel systems in the United States. Labov, Ash, and Boberg (in press) recorded 700 individuals across the country as part of their telephone survey (TELSUR) project. Based on an acoustic analysis of the vowels contained in the utterances, they have mapped the major and minor regional dialects of American English. The resulting atlas provides evidence for the major vowel shift phenomena that are currently taking place in North American English, including the Northern Cities shift, the Southern shift, the low-back merger found in the west and upper midwest areas, and Canadian raising. In addition, Thomas (2001) used the individual vowel spaces of nearly 200 talkers in various locations around the country and of several ethnic backgrounds as the basis for his description of vocalic variation in North American English, including detailed discussions of the vowel systems of communities in Ohio, North Carolina, and Texas, as well as African American, Mexican American, and Native American varieties. Finally, many other researchers in the fields of sociolinguistics and dialect geography have conducted small-scale studies of the vowel systems of regions from Maine to Missouri to California. The combined result of this effort is the mounting evidence for an enormous amount of variation in speech production as a result of regional and ethnic boundaries.

Despite the obvious relationship between speech perception research and sociolinguistic research on variation in production, speech perception researchers and sociolinguists have been working in almost complete isolation from one another. Speech researchers are interested in discovering ways to understand and model how humans perceive, process, and encode language and are faced with questions about acoustic-phonetic invariance in the speech signal and the role of different types of variability in language processing. In addition, theoretical linguists have also been working under the assumption that language is a symbolic system with relatively fixed underlying representations. Variation at the phonetic level has not been considered relevant to understanding, modeling, or describing language under this symbolic view. Therefore, until recently, variation in speech was treated as a source of noise; that is, as an undesirable set of attributes that needed to be reduced or eliminated.

On the other hand, sociolinguists are interested in describing natural variation as it occurs on social, regional, and ethnic levels and are faced with questions about the social implications of variability such as stereotypes, prejudice, and language attitudes as they impact the classroom and the workplace. Until recently, however, the question of how variation in language is perceived, processed, and encoded by listeners in order to allow them to make social judgments based on speech had been largely ignored by both speech researchers and sociolinguists. In this paper, we discuss some of the progress that has been made over the last 15 years in addressing the relationship between speech perception and dialect variation, as well as the implications of this research for studies of human speech perception, speech recognition and synthesis technologies, and linguistic theory.

Where Speech Perception and Sociolinguistics Intersect

Researchers in the fields of sociolinguistics and speech perception have provided large amounts of evidence to support the notion that linguistic variation between talkers due to regional and ethnic differences is real and robust. What we know less about, however, is what naïve listeners know about these sources of variation. Sociolinguists have spent much of their time documenting the linguistic variation that exists (Labov et al., in press) and speech perception researchers have spent their time trying to reduce or eliminate the variation (or ignoring it altogether) (Johnson & Mullennix, 1996).

There are a handful of methodologies, however, that have been used to investigate the question of what naïve listeners know about ethnic and regional linguistic variation. Some of these methodologies stem from the social psychology literature, such as attitude judgments and the matched-guise technique. Others have been developed in the field of perceptual dialectology, such as map-drawing tasks and dialect consciousness studies. Still others stem from the forensic linguistics literature, such as accent imitation and caricature. Finally, more recently a few researchers have employed methods developed in cognitive psychology to explore the perception of variation in discrimination, matching, identification, and categorization tasks.

Map Drawing. One of the more unique methodologies employed by sociolinguistic researchers interested in the mental representations of dialect variation is the map drawing task designed by Preston (1989). In this task, naïve participants are given a map of the United States (or Brazil or Japan) and are asked to draw and label the areas where “people speak differently.” The results of these studies reveal that the cognitive maps that these participants have of dialect variation do not correspond to the dialect maps that are drawn by sociolinguists and dialect geographers. In fact, while most undergraduates in the United States will identify some portion of the country as “South” and most can reliably identify New York City as having its own accent, composite maps of groups of participants invariably have one or more regions that are not labeled at all. That is, unlike dialectologists, naïve participants in these studies appear to believe that some regions of the United States are accent-free. In addition, Preston (1986) had students in Indiana, Hawaii, New York, and Michigan complete this map drawing task and he found that where the students were from had an effect on how they drew the maps. In particular, the participants tended to label more dialect regions in close geographic proximity to themselves than farther away. This finding suggests that naïve listeners are sensitive to the variation that they hear through personal experience with and exposure to people from areas surrounding their hometown or state.

While this kind of task reveals something about the mental representations that naïve listeners have about dialect variation, the task itself is based on judgments made from memory that may be highly biased and unreliable. The underlying assumption of the map drawing research is that the participants have some kind of full-formed mental representation of what they think the speech of a certain region sounds like. The results of these studies may not be able to provide a full understanding of speech perception or dialect perception, however, because they are based on measures of memory, not perception. In order to address issues of speech perception and dialect variation, researchers need to obtain some kind of behavioral response to actual spoken language stimuli. For example, participants could be given a map of the United States and after listening to a short sample of speech, they would be asked to indicate on the map all of the places that the talker might be from. Such a perceptual categorization task would reveal not only the participants’ perception of the speech sample under study, but also would provide information about how the participants mentally represent dialect regions, because they would be indicating on their map all of the places where they believe people talk in the same way as the talker in the stimulus item.

Attitude Judgments. In other research, Preston (1989) asked his participants to make judgments about the “correctness,” “pleasantness,” and intelligibility of the English spoken in each of the 50 states. In general, he found that participants rated their own speech as most intelligible and most pleasant, but he made their correctness ratings based on what seems to be a set of perceived notions about where Standard American English is spoken. Specifically, western and northern states were typically identified as having the most “correct” English by all participants, regardless of where they were from. Similarly, southern states were identified as having the least “correct” English, even by participants from southern Indiana, who speak a variety of southern American English. These findings reflect what Preston calls “linguistic insecurity.” Participants who are linguistically secure with respect to the variety of English that they speak are more likely to label their own variety as “correct” than participants who are linguistically insecure.

Like the map drawing task above, however, these attitude judgments rely on participant reports that are based on mental representations of language and there is no evidence to suggest that the participants necessarily have personal experience with or first-hand knowledge of the varieties of English that they label as least pleasant or most correct. These judgments could instead be highly biased and based on social stereotypes found in the media or perceived norms taught in the classroom by prescriptive grammarians.

Matched-Guise Technique. Another methodology that is commonly used in studies of language attitudes, particularly with respect to ethnic and racial varieties, is the matched-guise technique (Lambert, Hodgson, Gardner, & Fillenbaum, 1960). In a matched-guise experiment, listeners hear utterances read by a single talker assuming multiple guises (e.g., dialects, varieties, or languages) and are asked to rate the talker on scales such as intelligence, friendliness, and socioeconomic status. By controlling the voice qualities of the talker by using only a single talker, researchers can be more confident that their results point to attitudes toward phonological properties of language varieties and not to inherent differences in voice quality between talkers of different varieties. Studies of this kind often find that nonstandard language varieties are rated lower than standard varieties on scales related to "intelligence" by all listeners, revealing a general tendency to relate linguistic standardness with intelligence (Linn & Pichè, 1982; Luhman, 1990). However, it is also often the case that speakers of nonstandard varieties will rate those varieties highly on scales related to "friendliness," showing solidarity with speakers of the same variety (Linn & Pichè, 1982; Luhman, 1990). These types of studies suggest that listeners can and do make a number of attitudinal judgments about a talker based on his or her speech and that in many cases, these judgments correspond to social stereotypes or prejudices often associated with the group that is represented by a certain language variety.

In these sociolinguistic and social psychology studies, there is no way to separate the attitude judgments made by the listeners from their ability to recognize the dialect of the speaker. The analysis of results collected using the matched-guise technique often assumes that the listeners first identified the racial, ethnic, or regional accent of the talker before making their attitudinal response. However, listeners in these tasks are rarely asked to identify where the speaker is from before (or after) making their ratings. It therefore seems premature to conclude from these studies that listeners think that speakers of Appalachian English, for example, are friendly and unintelligent when in fact the only conclusion that can be drawn is that when the talker is speaking in an Appalachian English guise, the listeners rate him or her as being friendly and unintelligent (Luhman, 1990). In addition, the issue of native-like performance in all of the guises used in this kind of study is often overlooked. The crucial assumption made in this research is that the talker is equally competent in all of the guises he or she uses. It is difficult to know to what extent the talker truly controls each dialect and to what extent the characteristic or stereotyped features of each dialect are merely caricatured.

Caricatures. A similar method that has been used in the forensic linguistics literature but that may also provide insight into what people know about language varieties is an imitation or caricaturization task. In one such study, Markham (1999) asked eight native speakers of Swedish to read a prepared passage and an unfamiliar passage using a number of different regional accents. He then asked linguistically-trained judges to listen to the passages and identify the accent as well as rate the reading on its naturalness and purity. Markham found that some talkers were indeed able to convincingly imitate some accents, even for native listeners of that accent. These results suggest that in some cases, listeners can not only perceive and represent the variation in the language around them, but they can also reproduce the phonological characteristics of non-native varieties accurately. By including both a prepared passage and a sight-reading passage, Markham was able to elicit several levels of dialect imitation productions.

An interesting follow-up and extension to this study would be to play the speech samples to untrained native listeners of the different varieties represented and ask them to identify where the talkers were from and rate the nativeness or naturalness of the productions. This kind of follow-up would permit an examination of what the naïve listener knows about his or her own variety, as well as provide another measure of the talkers' abilities to imitate and reproduce non-native varieties.

Dialect Consciousness. In a slightly different approach to the investigation of what listeners know about the linguistic features of varieties of their native language, Mase (1999) conducted what he called a "dialect consciousness" study. He asked a group of Japanese participants to list characteristics of Japanese dialects that they perceived as being different from their own. The participants were able to provide grammatical, phonological, and lexical differences that distinguished their own dialect from the speech of the region in question. In addition, the features that the participants listed were typically those which are unique to a given region, and not those that are found in multiple dialects. That is, the participants were sensitive to the features that were characteristic of a single dialect as opposed to features that defined a broader region or group of dialects. Mase also studied the varieties actually spoken in the regions about which he had collected dialect consciousness data. He found that the characteristics provided by his participants were in large part quite accurate, although some of the properties tended to be older features that were used predominantly by the oldest generation or had died out completely, revealing a tendency for participants to report stereotypes that no longer reflected reality.

A parallel study has not been carried out in the United States, but a comparison of the results of the map-drawing task from American and Japanese participants suggests that Japanese speakers appear to have a better sense of the regional language varieties spoken in Japan than Americans do of American English varieties. It would be interesting to see to what degree native speakers of American English could correctly identify characteristics of southern speech or of a Boston accent. By investigating what features people think characterize a dialect, researchers may gain some insight into what features they are paying attention to when trying to determine where a person is from.

Phonological Description. Sociolinguists may not have spent their time asking naïve listeners to describe characteristic features of language varieties in the United States, but researchers such as Labov have devoted a great deal of effort to the description of variation in this country. As mentioned earlier, Labov et al. (in press) have compiled a large corpus of spoken language over the telephone and have analyzed the vowel productions of 700 speakers of American English. The results reveal several vowel shifts in progress. These data also provide a coherent account of the variation and variability in vowel production in the United States. While Labov's research is interesting and important from a documentation standpoint and while it provides researchers who are interested in variation with an excellent starting point for examining variation on a smaller scale in individual regions, states, or cities, the methodology is somewhat limited because it does not provide information about how this variation is actually perceived by naïve listeners.

Vowel Matching. One technique that does assess naïve listeners' perception of variation in production is the vowel-matching task used by Niedzielski (1999) in her study of the perception of the Northern Cities shift in Detroit English. In this task, listeners heard sentence-length utterances and were asked to select one of six synthesized vowel tokens that they thought matched the vowel in the target word in each sentence. Half of the listeners were told that the talker was from Detroit (as she actually was) and half of the listeners were told that the talker was from neighboring Canada. Niedzielski found that listeners who were told that the talker was from Canada most often selected the synthetic token that matched the actual vowel as the "best match." However, the listeners who were told that the talker was from Michigan most often selected the synthetic token that corresponded to a canonical (i.e., unshifted) vowel as the "best match." These results suggest that vowel perception is mediated by "knowledge" about the talker, such as where the listener believes the talker is from.

Niedzielski's (1999) conclusion was that Detroiters perceive themselves as speaking "standard" English, but that they perceive Canadians as speaking "with an accent" and this affected their perception of the vowels that they heard. One problem with this interpretation, however, lies in the design of the task. The listeners in Niedzielski's study were told to select the "best match" from six synthesized vowel tokens as part of a project on improving speech synthesis. It is possible and very likely that the group of listeners who were told that the talker was from Detroit selected canonical vowels because they wanted to be "helpful" to the experimenter by selecting the "best" vowels and not the "best match" vowels. In addition, although synthesized speech can be useful in tasks like this in which a range of tokens that are carefully controlled for formant values is necessary, synthetic speech samples are less natural than human speech productions and therefore research relying on behavioral responses to synthetic speech should be supplemented by converging evidence from studies involving natural speech.

Using a number of different methodologies from a variety of subfields of linguistics and psychology, researchers have begun to collect evidence to support the proposal that people can and do perceive and encode the variability in the speech they hear around them. Map-drawing, attitude judgment, and matched-guise tasks can provide researchers with valuable information about how people conceptualize the varieties of their native language. Caricature and dialect consciousness studies provide information about what the salient properties of a given language variety are and, in the case of caricatures, provide some insight into how well people can translate the knowledge they gain about linguistic variability through perception to production. Phonological studies of linguistic variation provide researchers with a basis for discussing what naïve listeners do and do not know about variation through thorough linguistic description. Finally, vowel-matching and other similar paradigms in cognitive psychology allow researchers to investigate perception of variation at a lower level of representation than the other kinds of tasks because, in ideal situations, they do not require the listeners to make more complex attitude judgments about the talkers.

The map drawing task and its associated attitude judgment tasks provide useful information about how naïve participants represent dialect variation. They help researchers answer questions such as: Where do people think people sound like each other? How many accents of American English do people think there are? What associations do people have with certain regional varieties? What do the mental maps of dialect variation look like for non-linguists? The matched-guise technique also provides some useful information about the kinds of attributes that naïve listeners associate with certain speech patterns. This methodology helps researchers to answer questions such as: What kinds of judgments do people make about people who talk a certain way? What kinds of attitudes towards people of certain races, ethnicities, socioeconomic status, and regional backgrounds can be elicited from listeners based on speech?

Caricature and dialect consciousness studies have focused on a slightly different aspect of perception, related to issues about naïve listeners' awareness of linguistic features. These kinds of studies help researchers answer questions such as: What do naïve listeners know about the linguistic features of other varieties? How accurate are listeners in describing or imitating characteristic features of other varieties? Can they imitate those features? Are their imitations good enough to fool native listeners? While these are all interesting questions, they are focused on the higher level of attitude representation, rather than perception.

Phonological descriptions of language varieties provide answers to questions about what the actual characteristic features of a dialect are. These descriptions can also help researchers answer questions such as: Where do people speak the same? How many dialects are there of American English? Which features are shared by multiple dialect groups? Which features are unique to a single region, ethnicity, or social class? Finally, the vowel-matching task provides some information about lower-level perception of variation. Niedzielski's (1999) study provided some initial answers to questions such as: How is perception influenced by a listener's beliefs about the utterance? To what extent does information beyond what is available in the acoustic signal impact perception?

The questions that remain to be investigated in dialect perception are those questions related to how listeners can actually use information in the speech signal to identify where a talker is from. Related issues include questions such as: What kinds of information does a listener encode with respect to dialect variation in everyday language situations? How is this information used in speech perception and language processing? How does linguistic experience with talkers from a variety of dialects affect a listener's ability to discriminate, identify, or describe different language varieties?

These kinds of questions can be explored using a wide variety of techniques developed in cognitive psychology, cognitive science, speech perception and spoken word recognition research. There are numerous experimental paradigms available in the speech perception literature that will allow researchers to investigate the perception of variation. For example, studies of dialect recognition or categorization based on actual speech samples can provide new insights about what information about language variation is actually encoded in memory. Perceptual learning paradigms can be used to examine the role of linguistic experience in dialect identification, categorization, and discrimination. There has been some progress in this direction already over the last few years. The application of experimental methods to the study of linguistic variation should provide new insights into dialect perception and complement some of the earlier research using traditional sociolinguistic and social psychology methods.

Dialect Categorization

Dialect categorization studies are quite limited in the literature, but several researchers have developed methodologies to determine whether listeners can identify where a talker is from based only on a short speech sample. These perceptual studies employ traditional identification or categorization methodologies developed in the field of cognitive psychology for studying speech perception and spoken word recognition. Listeners hear short segments of speech spoken by a number of talkers and are simply asked to identify where they think the talker is from, using either a closed-set categorization task or an identification task. While these kinds of studies cannot answer questions about how the listeners use their knowledge of variation to make judgments about the talkers, they can provide new information about how listeners can use their knowledge of variation to determine where the talker is from. In combination with acoustic analyses of the speech signal and/or synthetic manipulation of the speech to highlight certain features, these kinds of studies can also be used to answer basic questions about which acoustic properties of the speech signal are most salient to listeners in identifying a talker's dialect. Through the study of relevant cues in dialect categorization, we can better determine what kinds of information about dialect variation are encoded, stored, and represented by the naïve listener based on his or her everyday experiences with linguistic variation in the environment.

Purnell, Idsardi, and Baugh (1999) conducted an implicit dialect identification experiment using the matched-guise technique. A single male talker using three racial guises (African American Vernacular English, Chicano English, and Standard American English) left answering machine messages for landlords in five neighborhoods in the San Francisco area. The researchers measured dialect identification by examining the relationship between the number of returned phone calls leading to appointments with a landlord from each neighborhood and the minority population living in each neighborhood. They found that the number of appointments for the Standard American English guise remained relatively constant across all five neighborhoods. However, the number of appointments for the African American Vernacular English and the Chicano English guises declined with the population of minorities in the neighborhood. Purnell et al. concluded that the landlords could identify the dialect, and therefore race, of the talker from just a brief sample of speech left on an answering machine.

Baugh (2000) has described the behavior of the landlords as "linguistic profiling" and has appeared on National Public Radio to discuss the findings of his study. While the issues related to racial identification are important, the original study itself was fundamentally flawed in several ways. The first flaw has to do with the matched-guise technique itself. As mentioned above, there are serious concerns about the ability of a single talker to produce utterances natively in multiple guises. The talker in the Purnell et al. (1999) study may not have had equally good control of all three guises. Second, the authors acknowledged that the dialects they used were "broadly" defined, but it is well known in the linguistic literature that variation among white speakers is much more regionally based than variation among African American speakers. Therefore, it is possible that similar results could be obtained using a northern white guise, a southern white guise, and a New York City guise. A relationship then might become apparent between perceived socioeconomic status and number of appointments made by the landlords that also corresponds to the mean socioeconomic status of a given neighborhood. While the results of this study do seem to show that people use their perception of dialect in making decisions in everyday life, the experiment itself did not control for the relationship between dialect and socioeconomic status.

In a more explicit study of dialect identification, Preston (1993) asked undergraduates in Michigan and Indiana to identify nine male talkers on a north-south continuum between Dothan, AL and Saginaw, MI. The talkers were all middle-age males and the speech samples were short utterances taken from longer narratives. The listeners heard each talker only once and were asked to identify which of the nine cities they thought he was from. While listeners were quite poor at identifying exactly where each talker was from, they were able to distinguish between north and south. The major boundary for the two groups of listeners was slightly different, suggesting that dialect identification is partly based on where the listener is from.

More recently, Preston (2002) has suggested that the difference in the location of the north-south boundary for the two listener groups could be related to differences in what they were listening for. In particular, his other studies have shown that Michiganders pride themselves on having the most "correct" English in the United States, while Hoosiers pride themselves on sounding "pleasant." Preston suggested that one possible explanation for the difference in perceived boundary in the identification task is that the Michiganders were making their identifications based on "correctness," while the Hoosiers were making their identifications based on "pleasantness."

One weakness of this study is that the listeners heard each talker only once and had to assign one talker to each city. Listeners therefore had to make their first response without reference to anything other than their own speech. They could make the remaining responses by comparing the voice on that trial with all of the voices they had heard previously. It has long been known in social and cognitive pscychology that behavioral responses to stimuli require reference and comparison to a standard. If a benchmark is not provided by the experimenter, then the participant must rely on his or her own internal standard which may shift in the course of the experiment (Helson, 1948). In order to reduce the effects of shifts in participants' standards for comparison, an alternative to this study might provide listeners with all nine talkers and the option to listen to each one as many times as they want and in any order that they want so that the listeners could each create their own continuum of the nine talkers, without being restricted to a single repetition of each talker presented in random order. Despite this methodological problem, Preston's (1993) study provides some additional evidence that naïve listeners can distinguish northern talkers from southern talkers. This research also gives some insight into what listeners might be doing to make these judgments, but we still do not know what specific acoustic properties of the speech signal listeners are basing their "correctness" or "pleasantness" judgments on.

The first study that explicitly investigated dialect categorization was conducted by Williams, Garrett, and Coupland (1999) on varieties of English spoken in Wales. They recorded two adolescent males from each of six regions in Wales and two speakers of Received Pronunciation (RP) telling personal narratives. The authors played short segments of these narratives back to different groups of adolescent boys from each of the six regions and asked them to categorize each talker into one of eight categories (the six regions of Wales, RP, or "don't know").

Overall, the listeners were able to correctly categorize the talkers with about 30% accuracy. Williams et al. (1999) also looked at the performance of each group of listeners on the two talkers from their own region and found that performance on same-dialect talkers was not much better than categorization performance overall. The average performance was about 45% correct on talkers from the same region as the listeners. While the talkers were selected from a larger set of recordings based on phonological criteria established by the authors, there was a significant difference in how well the two talkers from any given region were identified by the listeners (from the same region or from a different region). The authors suggested that this difference may be due to the availability of more or fewer salient phonological cues in some narratives or to the content of the narratives themselves as revealing something about the region in which the talker lived. While this study used spontaneous speech samples as its stimuli with the expectation of revealing the "true" dialect of the talkers, the lack of segmental and contextual control of the stimulus materials themselves does not allow us to consider what the perceptual differences between the talkers should be attributed to.

Some recent work by Clopper and Pisoni has also focused on the question of dialect categorization. In one set of studies, we considered the question of how well listeners could identify where talkers were from and what acoustic-phonetic properties of the speech signal the listeners might be using to categorize the talkers (Clopper & Pisoni, submitted). We selected sentence-length utterances from eleven male talkers in their twenties from each of six dialect regions in the United States from the TIMIT Acoustic-Phonetic Continuous Speech Corpus (Zue, Seneff, & Glass, 1990). Participants listened to the sentences and were then asked to categorize each talker into one of the six regions. In the first two phases of the experiment, listeners heard each of the talkers reading the same sentence. In the final phase of the experiment, listeners heard each of the talkers reading a different novel sentence.

Like Williams et al. (1999), we found that our listeners were only about 30% accurate in categorizing the talkers. Figure 1 shows the overall performance of the listeners in each of the three phases of the experiment. A clustering analysis on the confusion matrices of their responses revealed that listeners were not randomly guessing, but instead that they were making broad distinctions between New England, southern, and western talkers. As an example, the clustering solution for the sentence, "She had your dark suit in greasy wash water all year" is shown in Figure 2. In this representation of perceptual similarity, perceptual distance is represented by the lengths of the vertical branches. It should be noted that the three perceptual clusters roughly correspond to the three major regional dialects of American English that Labov and his colleagues have discussed in the phonological variation literature (Labov, 1998). Therefore, while overall performance was just above chance in terms of categorization accuracy, the results of the clustering analysis suggest that the listeners were responding in a systematic fashion and made categorization judgments based on three broader dialect clusters than those presented as response alternatives.

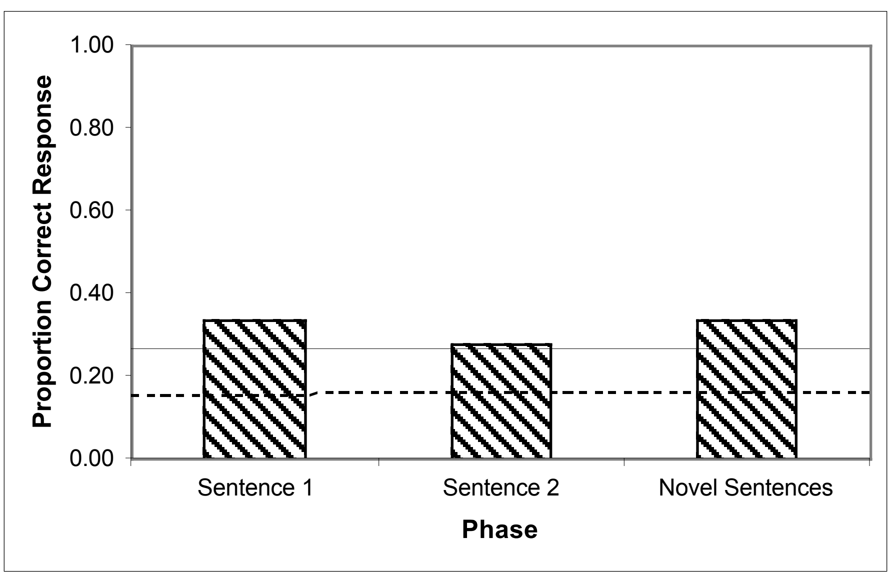


Figure 1. Proportion correct responses in each phase of the dialect categorization task. Chance performance (17%) is indicated by the dashed line. Performance statistically above chance (25%) is indicated by the solid line.

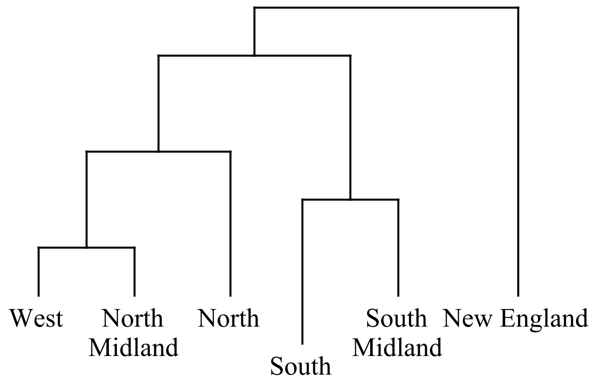


Figure 2. Clustering solution for the sentence, 'She had your dark suit in greasy wash water all year.' Perceptual distance is represented by the length of the vertical branches.

The sentence materials used in the categorization study were also subjected to an acoustic analysis. The acoustic measures confirmed that the talkers could be differentiated in terms of their dialect based on a number of reliable, well-defined acoustic-phonetic properties. In a logistic regression analysis, we found that there were seven acoustic-phonetic cues that were good predictors of dialect affiliation for our talkers. Table 1 shows the significant regression coefficients for the multiple regression analyses on the acoustic measures and dialect affiliation.

	Significant Variables	Regression Coefficients	Overall r^2
New England	r-fullness	-.01	.33
	/æ/ backness	.02	
North	/oʊ/ diphthong	-.01	.21
	/æ/ diphthong	-.01	
North Midland	n/a		
South Midland	/u/ backness	-.01	.19
	/oʊ/ backness	.01	
South	fricative voicing	3.4	.21
West	n/a		

Table 1. Results of the logistic multiple regression analysis on acoustic-phonetic properties and talker dialect affiliation. For each of the dialect groups, the significant acoustic measures are shown with their regression coefficients and the overall r^2 showing model fit.

A similar regression analysis of the results of the categorization study with the results of the acoustic analysis revealed that listeners were attending to only four of the seven available cues in the speech signal. They were also attending to an additional 12 cues that were not good predictors of the dialect affiliation of these talkers. The results of this second set of analyses are shown in Table 2. The four overlapping cues revealed listeners' sensitivity to stereotypes (New England r-lessness and North /oʊ/ pronunciation) and to prominent but less stereotyped variations (New England /æ/ backing and South Midland /u/ fronting). The results of the categorization study and the acoustic analysis together suggest that listeners can broadly categorize talkers by dialect and that they are able to make use of several reliable and robust cues in the speech signal to do so.

	Significant Variables	Regression Coefficients	Overall r^2
New England	r-fullness	-.36	.39
	/æ/ backness	.34	
	/oʊ/ diphthong	-.22	
	vowel brightness	.21	
North	/oʊ/ diphthong	-.38	.27
	/u/ backness	.29	
North Midland	/oɪ/ diphthong	.56	.31
South Midland	/u/ backness	-.26	.38
	vowel brightness	-.34	
	fricative voicing	.33	
South	/oɪ/ diphthong	-.39	.49
	/oʊ/ diphthong	.33	
	/u/ backness	-.33	
	/oʊ/ backness	.31	
	/æ/ diphthong	.20	
West	/oɪ/ diphthong	.40	.16

Table 2. Results of the linear multiple regression analysis on the acoustic-phonetic properties and perceptual categorization responses. For each of the dialect groups, the significant acoustic measures are shown with their regression coefficients and the overall r^2 showing model fit.

In a follow-up to the dialect categorization study, we investigated the role of the residential history of the listener on dialect categorization performance. In several studies, Preston (1989; 1993) has shown that participants from different parts of the country perform differently on his map-drawing and attitude judgment tasks. In our study, we asked two groups of listeners to carry out the same dialect categorization task described above. The first group ("homebodies") consisted entirely of listeners who had lived exclusively in Indiana. The second group ("army brats") consisted entirely of listeners who had lived in at least three states (including Indiana). We hypothesized that the listeners in the "army brat" group would perform better on the categorization task than the "homebodies" because through their real-life experiences living in a number of different places they would have been exposed to more variation than listeners who had lived exclusively in only one state.

Our results confirmed this hypothesis. The listeners in the "army brat" group performed better overall than the listeners in the "homebody" group. Figure 3 shows the proportion correct performance in each of the three phases for each of the listener groups. The clustering analysis on the data in this experiment also revealed differences in the perceptual similarity spaces of the dialects for the two listener groups, although the overall finding for both groups reflects the basic three-cluster structure (New England, South, West) found in the first experiment. These results confirm our intuition that personal experience with linguistic variation is an important contributing factor in how well people can identify where talkers are from based on their speech.

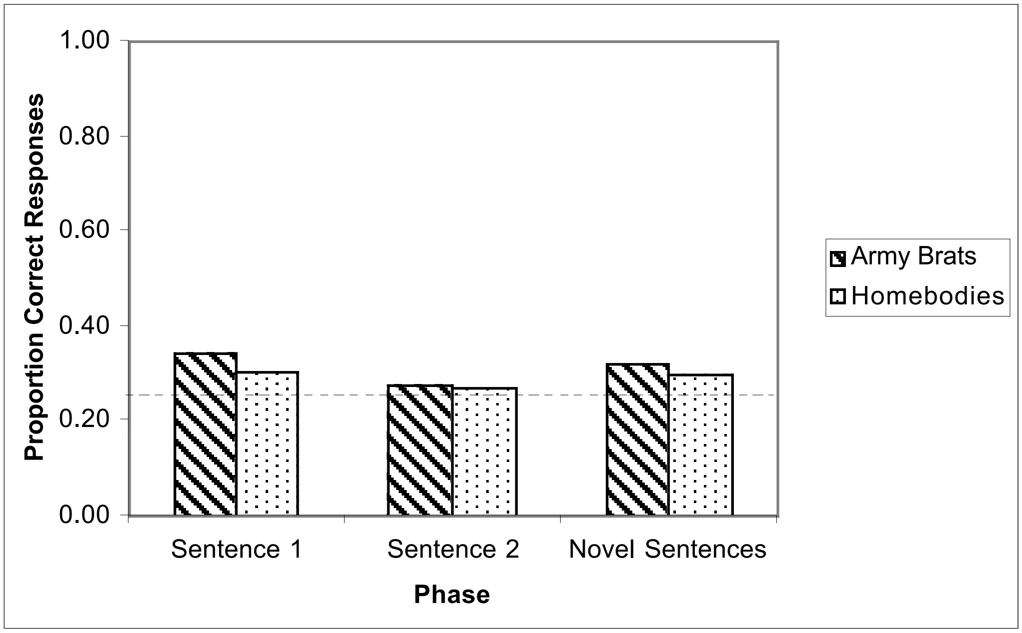


Figure 3. Proportion correct responses in each phase of the dialect categorization task for the Army Brat listener group and the Homebodies listener group. Chance performance (17%) is indicated by the dashed line. Performance statistically above chance (25%) is indicated by the solid line.

Perceptual Learning of Dialect Variation. Training and perceptual learning studies are often used in cognitive psychology to ensure that poor performance on a given task is not due merely to the participants' unfamiliarity with the task itself and to determine how much participants can improve and at what level their performance will asymptote (Green & Swets, 1966). Therefore, in order to determine whether or not personal experience in a laboratory setting would produce improvements in categorization performance, we conducted a set of short-term perceptual learning studies in which listeners were asked to learn to categorize a subset of the talkers used in the previous categorization tasks and then to generalize to new talkers. One group of listeners was trained to identify one talker from each of the six regions (the "one-talker" group). A second group of listeners was trained to identify three talkers from each of the regions (the "three-talker" group). Training consisted of three phases in which both groups of listeners heard sentences and were asked to categorize the talker by dialect. In the first two phases, the talkers all read the same sentence. In the third phase of training, every talker read a different, novel sentence. Feedback was given after every trial to aid in learning. Following the three training phases, the listeners participated in a test phase using the same talkers as in the training phases but without feedback to ensure that they had learned which talkers were from where. The last phase of the experiment was the generalization phase in which the listeners heard sentences read by all new talkers and were asked to categorize them without feedback. In both the test and generalization phases, the talkers all read different, novel sentences.

Categorization performance results are shown in Figure 4 for each of the five phases of the experiment for each of the two groups of listeners. While the one-talker group performed better in the training phases of the experiment, the three-talker group performed better in the generalization phase. This cross-over effect suggests that exposure to greater variation in training may produce more difficult initial learning in the training phases, but it results in better generalization to new talkers at test. Despite the fact that the training sessions for both groups were relatively short in comparison to other types of language-based perceptual learning experiments, listeners in the three-talker group were better able to categorize new talkers than listeners in the one-talker group. These results on perceptual learning of dialect variation suggest that even when explicit instructions are not given about how to do the task, listeners know what to listen for and can extract information out of the acoustic signal that helps them in identifying the dialect of other unfamiliar talkers with very little exposure to the stimuli.

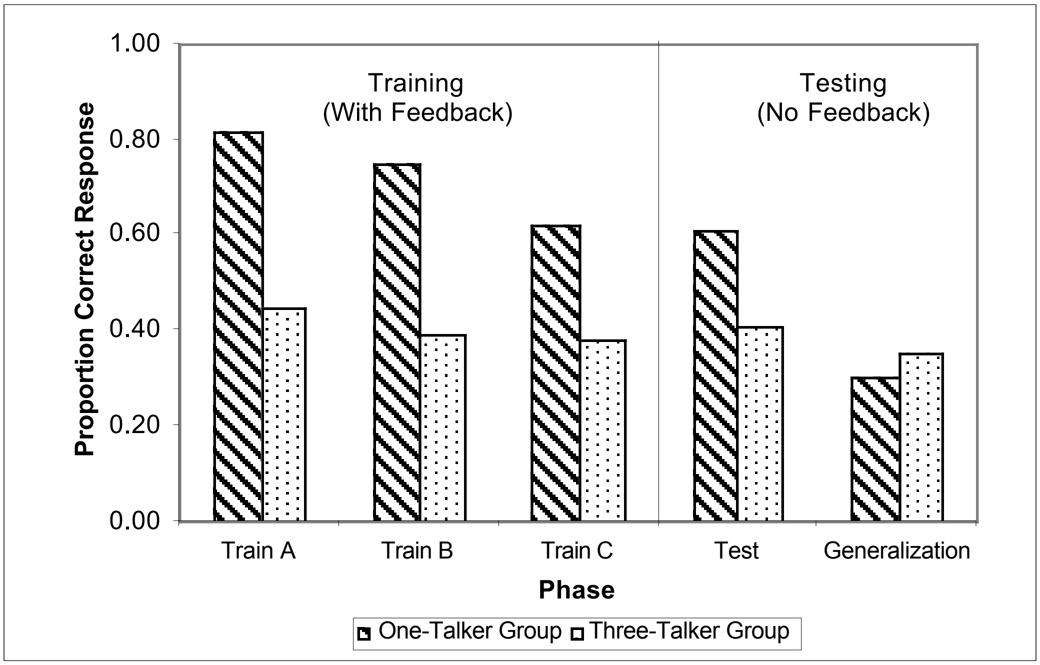


Figure 4. Proportion correct responses in each phase of the perceptual learning experiment for each of the listener groups.

The dialect categorization studies discussed above have a variety of goals with respect to the theoretical issues they wish to address. The matched-guise task focused on the judgments and decisions participants made based on their perception of where the talker was from (Purnell et al., 1999). The dialect categorization studies focused on how listeners made judgments about where a talker was from and what acoustic-phonetic properties of the speech signal the listeners were using to make such identifications (Clopper & Pisoni, submitted; Preston, 1993; Williams et al., 1999). Finally, the perceptual learning study examined the role that experience and learning have on dialect categorization abilities in naïve listeners. Despite these disparate goals, however, there is one theoretical claim that the results of all of the studies make that cannot be ignored: variation matters in the perception of spoken language. Naïve listeners can make reliable judgments about where an unfamiliar talker is from without explicit instructions about what to listen for. This ability suggests that listeners retain some kind of mental representation in memory of the varieties of their native language and that these representations develop naturally through a person's experience with and exposure to his community and the world at large. Specifically, recent findings from our lab have shown that greater personal experience and exposure with multiple dialects leads to better performance on the dialect categorization task. In addition, experience with greater variability of stimuli in perceptual learning paradigms of dialect variation also leads to better categorization performance on unfamiliar talkers. Therefore, experience both in real life and in the laboratory contribute to the information that listeners encode about the variation that they hear in the language around them.

Looking Forward

The relatively small literature investigating the relationship between dialect variation and speech perception in the laboratory means that there is still much work to be done before we can truly claim to understand how dialect variation is perceived, encoded, and represented in memory by naïve listeners. What little research has been done suggests that methodologies such as categorization tasks and perceptual learning tasks from the cognitive psychological literature and new methodologies developed by perceptual dialectologists such as map drawing tasks and the elicitation of dialect characteristics, as well as acoustic-phonetic analyses can provide converging information that will help us begin to answer fundamental questions about how listeners identify the dialect of a talker and how they use such categorizations once they have made them.

Methodological Extensions. There are several possibilities for extensions to the basic methodology of the dialect identification and categorization tasks. First, all of the major studies discussed above used only male talkers. This literature must be expanded to include studies of female speech and also studies in which the gender of the talkers varies from trial to trial. Sociolinguists have argued that women tend to be more conservative in their speech, often using fewer stigmatized forms (Labov, 1990). Speech stimuli recorded from male talkers might therefore be expected to reveal more regional or substratal forms. However, sociolinguists have also shown that women tend to be ahead of men in language changes in progress, regardless of whether the changes are above or below the level of conscious social awareness (Labov, 1990; Milroy & Milroy, 1993). Speech stimuli recorded from female talkers might therefore be expected to reveal current changes in progress. Niedzielski's (1999) perception work on Detroit speech suggests that listeners are not always aware of linguistic changes in progress, whereas Mase's (1999) dialect consciousness work suggests that people are aware of stigmatized forms that have become ingrained in popular social culture in the form of stereotypes. In order to understand the perception of both stigmatized and innovative forms, we need to extend the present research on dialect categorization to include responses to both male and female speech samples.

In addition, acoustic-phonetic research has traditionally involved only male talkers, due to the relative ease with which male formants can be measured as compared to female or child formants. However, recent acoustic analyses of male and female speech have shown that there are important acoustic differences between male and female speech in terms of segmental reductions (Byrd, 1994) and voice quality (Klatt & Klatt, 1990). Byrd's (1994) work also suggested an interaction between regional dialect and gender in segmental reduction that provides further support for the need to extend dialect categorization research to female talkers.

Second, the relatively poor performance of the listeners in the categorization tasks, their apparent ability to make broad categorical distinctions, and Preston's (1986; 1989) findings that naïve participants do not have cognitive maps that correspond to linguists' maps of dialect variation all suggest that in conducting these categorization tasks, we might want to reconsider the response format and alternatives that we provide for our listeners. Perhaps fewer response alternatives, representing the broad categories they show repeatedly in these tasks (e.g., North, South, and West) would result in better performance because it more directly reflects how listeners perceive and represent linguistic variation. Another alternative to the multiple choice tasks used in the studies discussed above would be a simple binary forced-choice distinction task in which listeners have to respond whether or not the talker has the same dialect as they do (e.g., "sounds like me" or "does not sound like me").

Third, all of the results described in this paper have relied on accuracy data in behavioral tasks. However, another common dependent measure used by psychologists interested in perception and processing is response latency. By adjusting the methodology slightly to force listeners to respond under time pressure, one could elicit response latency data that might also provide some insight into the underlying process of how listeners make their decisions. Are some varieties easier (faster) to identify than others? Are some listeners faster to respond than others? Do listeners respond faster to talkers from their own region than to talkers from other regions? These are all questions that have not been investigated previously but that might provide further evidence about the role of variation in language processing, perception, and encoding.

The perceptual learning study discussed above also represents merely the tip of the iceberg of possible training methodologies that could be employed to investigate how listeners learn what to attend to. The one study that has been conducted so far involved short-term learning in a single session of less than one hour and short-term retention with the generalization phase immediately following the last training phase. While there was a small amount of improvement for the group who was trained on more talkers over the group who was trained on fewer talkers, neither group performed much above the levels of untrained listeners in other experiments. The question then remains at what level of performance the listeners would asymptote, if the training were continued over a number of sessions over a number of weeks. Similarly, the question also remains as to how long listeners would be able to retain whatever they had learned in the training sessions. Would the listeners exposed to more talkers still perform better on novel talkers after one day or one week?

In addition, the explicitness of the instructions could also be manipulated to determine whether or not asking listeners to focus on certain things affects their ability to learn to categorize talkers by dialect. For example, would the instruction to "focus on the vowels" cause listeners to improve even more over those without specific instructions, or do their strategies already include such a focus? Finally, like the categorization experiments above, the materials in both training and generalization phases need to be more varied to include not only talkers of both genders but also other kinds of utterances such as syllables, words, sentences, and perhaps even longer passages of connected speech.

Listener Populations. Another similarity between all of the categorization studies discussed above is that the talkers and listeners were all young to middle-aged normal hearing adults. It would be useful to extend this research to populations such as infants, children, older adults, non-native speakers, and hearing-impaired children and adults to investigate the effects of age, language background, and hearing impairment on dialect categorization. In particular, studies with infants and children would allow us to determine at what age the abilities to discriminate and categorize dialects arises. We might expect this ability to arise quite early in development, given some of the findings in the infant and child speech perception literature. For example, Houston and Jusczyk (2000) found that 10.5-month-old infants could separate the linguistic content of the speech signal from the indexical properties of the talker better than 7.5 month olds. Spence, Rollins, and Jerger (2002) have shown that 3-, 4-, and 5-year olds can use indexical information to identify cartoon characters by their voice. These findings suggest that indexical properties such as dialect variation are encoded in speech perception early in development and that children quickly learn to separate these talker-specific properties from the linguistic meaning.

Dialect categorization studies with older adult listeners would add to the discussion of the role of linguistic experience in dialect categorization. One hypothesis might be that older adults would perform better than younger adults because they have had more time to come into contact with more variation. In their study of dialect categorization in Wales, Williams et al. (1999) used two populations of listeners, adolescents and schoolteachers. Although they did not provide a statistical comparison of the performance between the two groups, the schoolteachers performed better (52% correct) than the adolescents (30% correct). Williams et al. concluded from these results that linguistic experience comes with age and that the difference between the two populations could be attributed to the greater experience of the teachers with linguistic variation. If performance continues to increase with age and experience, we might expect to find better categorization performance for older adults than for the college-aged listeners used in most of the studies discussed above.

When it comes to dialect categorization by non-native listeners, we might be inclined to predict that they would perform more poorly on a categorization task than native listeners. First, non-native listeners would typically have less experience with and exposure to the variation in the target language.

Second, they might be less sensitive to the variation in a second language than native speakers, particularly with respect to phonetic variation within a single phonological category. However, Bradlow and Pisoni (1999) found that non-native listeners were not more susceptible to talker variability effects in word recognition than native listeners, suggesting that some kinds of indexical variability have the same effects on all listeners. Dialect categorization research using non-native listeners would be an important contribution to our understanding of how linguistic variation is perceived and in what ways perception is constrained by language background.

Research involving hearing-impaired populations would provide evidence for the robustness of variation in cases where the signal is degraded. A case study conducted in our lab of a post-lingually deafened adult cochlear implant user on the perceptual learning task with training on a single talker from each region revealed poorer performance than normal hearing listeners in all of the training phases and performance at chance on the generalization phase. These results suggest that at least some of the information that is encoded by normal hearing listeners in this perceptual learning task is either not available to or not encoded by cochlear implant users. In addition, research on both adult and pediatric cochlear implant users has shown that they perform more slowly and less accurately on talker discrimination tasks than their normal hearing peers, suggesting that indexical information is not perceived and encoded in the same way for the two populations (Cleary, 2002; Kirk, Houston, Pisoni, Sprunger, & Kim-Lee, 2002). More research on these and other clinical populations would provide even further insight into the kinds of information that are available to and encoded by listeners in making these kinds of categorization judgments about language variation.

Other Measures of Perception. Another approach to the study of the perception of variation that has barely been examined is the notion of perceptual similarity spaces. The clustering analyses that we conducted on the confusion matrices from our dialect categorization studies reflect just one method of determining the perceptual similarity of the dialects we studied. Our results suggested that the perceptual similarity between dialects is based in part on the phonological similarity of the dialects, but that it also might be influenced by stereotyped uniqueness of a given variety. In particular, New England and South were often the most distinct dialects for our listeners and these two dialects are both associated with a number of stereotyped features. Other methodologies from the cognitive psychology literature such as paired comparisons, free classification, and similarity ratings tasks would provide converging evidence for the similarities between dialects (and between talkers within a given dialect) as they are perceived by naïve listeners.

While there is an extensive literature that outlines acoustic-phonetic differences between dialects in terms of vowel production, a complete discussion of the perception of variation in speech would also require a prosodic analysis of different language varieties. For example, are there consistent differences between dialects in terms of speaking rate (e.g., do southerners really talk more slowly?), fundamental frequency modulation, and stress? In addition to measuring these differences, it might be informative for researchers interested in the perception of dialect variation in the United States to conduct a dialect consciousness study like Mase's with American English speakers. A speaker's ability to articulate just what the characteristic aspects are of a given variety (his own or another's) will certainly reflect at least in part how he has represented that variety in memory. The results of these and many other possible studies will lead us to some better answers about how language variation is perceived, processed, stored, and used in human speech perception and spoken language processing.

Finally, electrophysiological and neuroimaging approaches to the study of the perception of dialect variation have barely been explored. Conrey (2001) reported the results of a vowel merger perception experiment in which she recorded reaction times in a cross-modal semantic priming task. She found that the behavioral reaction times in her study correlated with prior electrophysiology research on semantic priming. This correlation suggests that electrophysiology research on the perception of vowel mergers might also reveal interesting results that would provide further insights into the perception of variation. In addition, fMRI research has revealed some cortical distinctions between how linguistic form and linguistic content are processed (Ni, Constable, Mencl, Pugh, Fulbright, Shaywitz, Shaywitz, Gore, & Shankweiler, 2000) and between first and second language processing in bilinguals (Kim, Relkin, Lee, & Hirsch, 1997). Future fMRI research on the processing of indexical variation and the role of linguistic experience with dialect variation in language processing may provide more information about how variation and variability are perceived, processed, and encoded by the human listener.

Implications for Speech Research, Speech Technology, and Theoretical Linguistics

There are many reasons why we need to gain a better understanding of dialect variation and perception. In terms of human speech perception, the more we know about how variation and variability are perceived, the better we will be able to understand and model spoken language processing. Many current models of speech perception assume that variation is stripped off early in a process of normalization so that the meaningful content of the signal can be recognized (Pisoni, 1997). This assumption is central to the traditional abstractionist view of speech and language as symbolic systems, in which the variation is treated as irrelevant noise. However, in order to completely understand the process of human speech perception, we need to understand how the sources of variability described by Klatt (1989) are perceived and encoded along with the linguistic message of the utterance. Researchers have only recently begun to abandon the traditional symbolic view of language and to investigate the contributions of linguistic variability in human speech perception. For example, in addition to the recent findings of the "army brat" study reported above, there is also an extensive literature on the role of talker variability and talker-specific information in speech perception that suggests that indexical properties of the talker are perceived and encoded by listeners in everyday linguistic tasks (e.g., Mullennix, Pisoni, & Martin, 1989; Nygaard, Sommers, & Pisoni, 1994). Dialect variation is clearly one of the indexical properties that is perceived and encoded in everyday language situations and its impact on speech perception deserves further investigation.

The implications for automatic speech recognition (ASR) systems with respect to variation and variability are perhaps even more striking. The variation and variability that exists in a single language is simply enormous and is constantly changing as the language changes. Human beings are able to adapt quickly to new talkers and linguistic changes, but ASR systems are still severely limited with respect to variation and change and require large amounts of training before they can accurately recognize speech. Ideally, ASR systems would be able to recognize not only a large number of lexical items, but also a large number of talkers and a large number of languages. However, most of the currently available commercial speech recognition systems are limited to a few talkers (e.g., personal computer speech-to-text software) or have limited vocabularies within a specialized domain (e.g., interactive automated flight information programs). One of the new areas of research in ASR systems is the "speech graffiti" project at Carnegie Mellon whose goal is to develop a universal speech interface that is more flexible than touchtone phone menu systems, but more rigid than a true natural language interface (Rosenfeld, Olsen, & Rudnicky, 2000). The idea behind the project is to build a human-machine speech interface that will be useful for an unlimited number of talkers across multiple domains, such as movie or apartment listings and flight information. The more we know about variation and how it is processed and encoded by human listeners, the more we will be able to apply our knowledge of human speech perception to building truly robust ASR systems.

Like ASR systems, speech synthesis technology is typically limited to a small number of voices and a limited vocabulary domain. The most natural synthetic speech can be built from the concatenation of resynthesized speech units smaller than the word, but larger than diphones. However, these systems are usually highly constrained in vocabulary. Successful speech synthesis of large vocabularies typically involves the concatenation of diphone strings, but the result is less natural speech (Black, 2002). Researchers at the University of Edinburgh in Scotland have been working to create a speech synthesizer using diphone concatenation that can produce speech in a number of different dialects of English, including Irish, Scottish, British, and American English varieties (Fitt & Isard, 1999). In addition to issues of prosody and sentence focus which remain problematic for speech synthesis programs (Wightman, Syrdal, Stemmer, Conkie, & Beutnagel, 2000), "natural" speech synthesis must also be able to replicate important human features of speaking style such as register shifts and dialect variation, given the importance of such factors in human communication and interaction (Giles & Bourhis, 1976). As we learn more about what parts of the acoustic signal are important for human listeners in identifying where someone is from, we will be better equipped to design synthetic speech production systems that exhibit the appropriate characteristics of a given dialect.

Finally, research on the perception of dialect variation also has some important implications for theoretical linguistics. Like many speech perception researchers, theoretical linguists typically assume that each lexical item specifies one underlying phonemic input that is transformed through serial derivation or parallel candidate selection into a phonetic output. Generative phonologists typically assume a one-to-one mapping between phonemic forms in the mental lexicon and phonetic outputs in production. However, results of the studies on dialect caricatures and dialect consciousness discussed above suggest that naïve listeners have multiple mappings between underlying and surface forms, both productively and conceptually. In the sociolinguistics literature, variable rule analysis has been adopted by many researchers to account for variable phonetic outputs given a single underlying form in a single talker (Labov, 1969). However, acknowledging and accounting for the possibility of a one-to-many relationship between phonemic representations and phonetic forms in production has yet to occur in the mainstream generative paradigm. The research discussed above, however, suggests that phonological variation is important in human speech perception and any model of phonology would be remiss in overlooking this physically and psychologically real aspect of human language.

Cognitive scientists have recently begun to embrace again the notion of embodiment and explore the relationship between cognition and human interaction with the world (Núñez & Freeman, 1999). Recent work in the fields of speech perception and sociolinguistics crucially reveals that language is more complex than a simple symbolic system and that the perception of speech involves not only extraction of the linguistic meaning of the utterance, but also a number of other processes including identification of some of the indexical properties of the talker. Language as a cognitive process is therefore embedded in our physical and social interaction with the environment and any viable model of language processing must account for the variability inherent in actual language use.

Researchers in the fields of social psychology, sociolinguistics, forensic linguistics, psycholinguistics, and cognitive psychology have all contributed to the growing literature on the relationship between regional, social, and ethnic language variation and speech perception. The results of this diverse set of studies reveal that naïve listeners are aware of linguistic variation to the extent that they can imitate it, describe it, and use it to identify where people are from and to make judgments about social characteristics of the talkers. The implications of this research are widespread as well, including issues related to models of human speech perception, speech perception in clinical populations, child language development, speech recognition and speech synthesis technologies, neural biology, cognition and language, and theoretical linguistics. There is much work still to be done in this area, however, as well as a need for multi-disciplinary discussion of the results of these many and varied studies and the implications of these results for our understanding of human language.

References

Baugh, J. (2000). Racial identity by speech. *American Speech, 75*, 362-364.

Black, A.W. (2002). Text-to-speech synthesis. Paper presented at the 143rd Meeting of the Acoustical Society of America Short Course on Speech Technology and Conversational Systems. Pittsburgh, Pennsylvania, June 1-2.

Bradlow, A.R. & Pisoni, D.B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America, 106*, 2074-2085.

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication, 15*, 39-54.
Cleary, M. (2002). Perception of talker similarity by normal-hearing children and hearing-impaired children with cochlear implants. Poster presented at the 143rd Meeting of the Acoustical Society of America. Pittsburgh, Pennsylvania, June 3-7.
Clopper, C.G. & Pisoni, D.B. (submitted). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics.*
Conrey, B. (2001). Effects of dialect on merger perception. Poster presented at New Ways of Analyzing Variation 30. Raleigh, North Carolina, October 11-13.
Fitt, S. & Isard, S. (1999). Synthesis of regional English using a keyword lexicon. *Proceedings of Eurospeech 1999*, 823-826.
Giles, H. & Bourhis, R.Y. (1976). Methodological issues in dialect perception: Some social psychological perspectives. *Anthropological Linguistics, 18*, 294-304.
Green, D.M. & Swets, J.A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *Journal of the Acoustical Society of America, 102*, 655-658.
Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychological Review*, *55*, 297-313.
Hillenbrand, J., Getty, L.A., Clark, M.J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099-3111.
Houston, D.M. & Jusczyk, P.W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1570-1582.
Johnson, K. & Mullennix, J. W. (Eds.). (1996). *Talker Variability in Speech Processing*. San Diego: Academic Press.
Kim, K.H.S., Relkin, N.R., Lee, K.-M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*, 171-174.
Kirk, K.I., Houston, D.M., Pisoni, D.B., Sprunger, A.B., & Kim-Lee, Y. (2002). Talker discrimination and spoken word recognition by adults with cochlear implants. Poster presented at the 25th Mid-Winter Meeting of the Association for Research in Otolaryngology. St. Petersburg, Florida, January 26-31.
Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 169-226). Cambridge, MA: MIT Press.
Klatt, D.H. & Klatt, L.C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87*, 820-857.
Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language, 45*, 715-762.
Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change, 2,* 205-254.
Labov, W. (1998). The three English dialects. In M. D. Linn (Ed.), *Handbook of Dialects and Language Variation* (pp. 39-81). San Diego: Academic Press.
Labov, W., Ash, S., & Boberg, C. (in press). *Atlas of North American English*. Mouton deGruyter.
Lambert, W., Hodgson, E.R., Gardner, R.C., & Fillenbaum, S. (1960). Evaluation reactions to spoken languages. *Journal of Abnormal and Social Psychology, 60*, 44-51.
Linn, M.D. & Pichè, G. (1982). Black and white adolescent and preadolescent attitudes toward Black English. *Research in the Teaching of English, 16*, 53-69.
Luhman, R. (1990). Appalachian English stereotypes: Language attitudes in Kentucky. *Language in Society, 19*, 331-348.
Markham, D. (1999). Listeners and disguised voices: The imitation and perception of dialectal accent. *Forensic Linguistics, 6*, 289-299.
Mase, Y. (1999). On dialect consciousness: Dialect characteristics given by speakers. In D. R. Preston (Ed.), *Handbook of Perceptual Dialectology* (pp. 101-113). Philadelphia: John Benjamins.
Milroy, J. & Milroy, L. (1993). Mechanisms of change in urban dialects: The role of class, social network and gender. *International Journal of Applied Linguistics, 3*, 57-77.

Mullennix, J.W., Pisoni, D.B., & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85*, 365-378.

Ni, W., Constable, R.T., Mencl, W.E., Pugh, K.R., Fulbright, R.K., Shaywitz, S.E., Shaywitz, B.A., Gore, J.C., & Shankweiler, D. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience, 12*, 120-133.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*, 62-85.

Núñez, R. & Freeman, W.J. (Eds.). (1999). *Reclaiming Cognition: The Primacy of Action, Intention, and Emotion*. Bowling Green, OH: Imprint Academic.

Nygaard, L. ., Sommers, M.S., & Pisoni, D.B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*, 42-46.

Peterson, G.E. & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175-184.

Pisoni, D.B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate, and perceptual learning. *Speech Communication, 13*, 109-125.

Pisoni, D.B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J.W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). San Diego: Academic Press.

Preston, D.R. (1986). Five visions of America. *Language in Society, 15*, 221-240.

Preston, D.R. (1989). *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics*. Providence, RI: Foris.

Preston, D.R. (1993). Folk dialectology. In D.R. Preston (Ed.), *American Dialect Research* (pp. 333-378). Philadelphia: John Benjamins.

Preston, D.R. (2002). The social interface in the perception and production of Japanese vowel devoicing: It's not just your brain that's connected to your ear. Paper presented at the 9th Biennial Rice University Symposium on Linguistics: Speech Perception in Context. Houston, TX, March 13-16.

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology, 18,* 10-30.

Rosenfeld, R., Olsen, D., & Rudnicky, A. (2000). A universal human-machine speech interface. *Technical Report CMU-CS-00-114*. Pittsburgh, PA: School of Computer Science, Carnegie Mellon University.

Spence, M.J., Rollins, P.R., & Jerger, S. (2002). Children's recognition of cartoon voices. *Journal of Speech, Language, and Hearing Research, 45*, 214-222.

Thomas, E. R. (2001). *An Acoustic Analysis of Vowel Variation in New World English*. Durham, NC: Duke University Press.

Wightman, C.W., Syrdal, A.K., Stemmer, G., Conkie, A., & Beutnagel, M. (2000). Perceptually based automatic prosody labeling and prosodically enriched unit selection improve concatenative text-to-speech synthesis. *Proceedings of the International Conference on Spoken Language Processing*.

Williams, A., Garrett, P., & Coupland, N. (1999). Dialect recognition. In D. R. Preston (Ed.), *Handbook of Perceptual Dialectology* (pp. 345-358). Philadelphia: John Benjamins.

Zue, V., Seneff, S., & Glass, J. (1990). Speech database development at MIT: TIMIT and beyond. *Speech Communication*, *9*, 351-356.

Speech Intelligibility of Children with Cochlear Implants and Children with Normal Hearing: A Preliminary Report[1]

Steven B. Chin and Patrick L. Tsai[2]

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

[1] This work was supported by grants from the National Institute on Deafness and Other Communication Disorders to Indiana University (R01DC00423 and T32DC00012). We are grateful to Susan T. Sehgal, Cara Lento Kaiser, Tara E. O'Neill, and Shelly P. Godar for assistance with data collection; to Theresa S. Kerr for database analysis; to Michael S. Vitevitch, Karen Iler Kirk, and Derek M. Houston for comments and suggestions; and to Luis R. Hernández and Jeffrey D. Karpicke for assistance with manuscript formatting. Finally, we very much appreciate the assistance and cooperation of the children who participated and their parents, as well as of the staff and administration of the Center for Young Children at Indiana University–Purdue University Indianapolis, in particular its director, Beth Jeglum.
[2] Department of Otolaryngology–Head and Neck Surgery, Indiana University School of Medicine (Chin) and Indiana University School of Medicine (Tsai).

Speech Intelligibility of Children with Cochlear Implants and Children with Normal Hearing: A Preliminary Report

Abstract. The objectives of the study reported here were (a) to assess the convergent validity of the Beginners' Intelligibility Test with respect to other measures of speech intelligibility in children, (b) to assess the development of speech intelligibility in children with normal hearing between the ages of 2 and 6 years, and (c) to compare the development of speech intelligibility in children with normal hearing and children with cochlear implants. The Beginners' Intelligibility Test, a sentence-repetition task, was administered to 50 children with normal hearing and 34 age-matched children who used cochlear implants. Responses were audiotape-recorded and presented to naïve adult listeners for transcription. Percent correct scores were compared for the effects of chronological age and hearing status. Consistent with previous studies of intelligibility, children with normal hearing achieved adult-like or near-adult-like speech intelligibility by around 4 years of age or shortly thereafter. Children with cochlear implants were considerably less intelligible than their chronological-age peers with normal hearing at all ages, although the intelligibility of children with cochlear implants did increase with chronological age through the latest age examined. Results from this study have important implications for the socialization and education of children with cochlear implants, particularly with respect to on-time placement in mainstream educational environments with same-age peers.

Introduction

Since the early 1990s, cochlear implantation has become a widely used treatment in cases of profound deafness in the pediatric population. Numerous studies have attested to the beneficial effects of cochlear implantation on the communicative abilities of children who use them. Insofar as spoken communication involves the transfer of information and knowledge, of paramount importance in assessing the effects of cochlear implantation is the connected speech intelligibility of the user of a cochlear implant, intelligibility referring here to "the degree to which the speaker's intended message is recovered by the listener" (Kent, Weismer, Kent, & Rosenbeck, 1989, p. 483) or "the comprehensibility of the specifically linguistic information encoded by a speaker's utterances" (Samar & Metz, 1991, p. 699).

A number of studies of the connected speech intelligibility of children who use cochlear implants dating from the early 1990s to approximately 1998 have been reviewed by Svirsky and Chin (2000), and the main results from previous research are (a) children's speech intelligibility improves from before implantation to after implantation and then further improves with increased use of a cochlear implant (e.g., Dawson et al., 1995; Mondain et al., 1997; Tobey, Angelette et al., 1991; Tobey & Hasenstab, 1991), and (b) cochlear implants support the development of speech production intelligibility at least as well as conventional hearing aids (depending on such factors as length of device use, age at device fitting, and amount of residual hearing; e.g., Miyamoto, Kirk, Robbins, Todd, & Riley, 1996; Miyamoto, Kirk et al., 1997; Miyamoto, Svirsky et al., 1997; Osberger, Maso, & Sam, 1993; Osberger, Robbins, Todd, Riley, & Miyamoto, 1994; Svirsky, 2000; Svirsky, Sloan, Caldwell, & Miyamoto, 2000). Other studies have examined the effect of communication mode on connected speech intelligibility (e.g., Tobey et al., 2000; Vieu et al., 1998) and the relationships between connected speech intelligibility and other communicative abilities (e.g., Chin, Finnegan, & Chung, 2001; O'Donoghue, Nikolopoulos, Archbold, & Tait, 1999). With the notable exception of several studies from University Hospital at the Queen's

Medical Center in Nottingham, United Kingdom (e.g., Allen, Nikolopoulos, Dyar, & O'Donoghue, 2001; Allen, Nikolopoulos, & O'Donoghue, 1998; O'Donoghue et al., 1999), which used rating scales, most studies of the intelligibility of children who use cochlear implants have used write-down (transcription) procedures. Speech materials used for the assessment of the connected speech intelligibility of children with cochlear implants have included sentences developed by McGarr (1983; e.g., Tobey et al., 1991); by Monsen (1983; e.g., Osberger et al., 1993); and by Osberger, Robbins, Todd, and Riley (1994). Materials from McGarr (1983) and Monsen (1983) were developed for assessing the intelligibility of children with hearing impairments and from Osberger et al. (1994) specifically for children who use cochlear implants (see also Svirsky & Chin, 2000; Svirsky, Chin, & Miyamoto, in press).

Previous research indicates that cochlear implantation for children supports the development of connected speech intelligibility at least as well as the use of conventional hearing aids. However, a variety of factors, including improved hardware and software, younger ages at implantation, and improvements in habilitation, have raised expectations to the point that language acquisition and development on a par with that of children with normal hearing is no longer considered unrealistic. With respect to connected speech intelligibility, however, it is not yet known how children who use implants compare to children who have normal hearing.

In comparing the connected speech intelligibility of children who use cochlear implants to that of children with normal hearing, one factor to consider is that although there is an obvious difference between children who use cochlear implants and children with normal hearing (the presence vs. absence of deafness), children with normal hearing do not in fact form a homogeneous population with respect to the production of speech and language. For the most part, research on connected speech intelligibility in children has been clinical, rather than theoretical. In addition to the extensive literature on the speech intelligibility of children with hearing impairments, the bulk of research on the speech intelligibility among children with normal hearing has been devoted to those children with such conditions as Down's syndrome (e.g., Chapman, Seung, Schwartz, & Kay-Raining Bird, 1998; Kumin, 1994), cerebral palsy (e.g., Clarke & Hoops, 1980), CNS lesions (e.g., Jerger, 1987), cleft palate (e.g., Keuning, Wineke, & Dejonckere, 1999), and autism (e.g., Koegel, Camarata, Koegel, Ben-Tall, & Smith, 1998). However, for children developing language in the absence of frank organically-based difficulties (including deafness), as recently as 2000, Gordon-Brannan & Hodson (2000) observed that "intelligibility data for young children with typical as well as disordered phonologies are generally lacking even though critical clinical decisions often depend on intelligibility" (p. 142).

The lack of definitive data can be traced in part to a lack of consensus regarding how to measure intelligibility and how to interpret the resulting data. In a review of evaluation procedures for intelligibility, Kent, Miolo, and Bloedel (1994) note that "although many would agree with Subtelney's (1977) comment that 'Intelligibility is considered the most practical single index to apply in assessing competence in oral communication' (p. 183), consensus withers when it comes to deciding how intelligibility should be measured and assessed" (p. 81). Additionally, most researchers regard connected speech intelligibility not as a monolithic sui generis factor, but rather as a factor dependent on, or at least related to, a variety of linguistic and extralinguistic characteristics. Weston and Shriberg (1992), for example, examined the influence of contextual and linguistic factors on intelligibility, finding that intelligibility outcomes were associated with utterance length, word position, the intelligibility of adjacent words, phonological complexity, grammatical form, and syllabic structure. Kent et al. (1994) further suggested that a number of factors affect speech intelligibility. Besides the basic competence of the speaker, these include the nature of the spoken material, the context of communication, the listener's familiarity with the speaker, contextual support for the message to be transmitted, clarity of the visual and acoustic signals of speech, and other environmental and linguistic factors.

In spite of the difficulties inherent in assessing connected speech intelligibility in young children and in interpreting results, a handful of studies have addressed the identification of benchmarks in the development of intelligibility in children. Weiss (1982), cited by Gordon-Brannan (1994), reported expectations for ranges of intelligibility for young children: 26% to 50% intelligible by age 2, 51% to 70% intelligible by age 2;6, 71% to 80% intelligible by age 3;0, 81% to 90% intelligible by age 3;6, and 100% intelligible by age 4;0. Gordon-Brannan (1994) points out, however, that the procedures for obtaining these data were not made clear.

In a study of 235 children, Coplan and Gleason (1988) asked parents to respond to the question "How clear is your child's speech? That is, how much of your child's speech can a stranger understand: (1) less than half, (2) about half, (3) three quarters, or (4) all or almost all?" Smoothed curves for age of emergence of 50%, 75%, and 100% intelligibility were derived, and cutoff ages were determined at which 90% of respondents ascribed each level of intelligibility to their child. The cutoff age for 50% intelligibility was 22 months, for 75% intelligibility 37 months, and for 100% intelligibility 47 months. As a comparison, Coplan and Gleason further cited anecdotal evidence from Weiss and Lillywhite (1976): 50% by age 2, 75% by age 3, and 100% by age 4.

Gordon-Brannan and Hodson (2000) examined intelligibility in 48 prekindergarten children with normal hearing ranging in age from 4;0 to 5;6 (mean = 4;7), dividing them into four groups based on the percentage of words transcribed correctly from a connected speech sample by unfamiliar listeners. Intelligibility scores for a "severe" group (the group with the lowest intelligibility scores) ranged from 16% to 63%. For the other three groups (moderate, mild, and adult-like), scores ranged from 68% to 100%, with a mean of 85%. Gordon-Brannan and Hodson suggested that for a child of 4 years or older, a score of less than 66% (2 standard deviations below the mean) might be a potential indicator of speech difficulty.

Despite variations in methodology, these studies appear to concur that children with normal hearing become fully intelligible by approximately age 4. The present study addresses two questions. First, does use of the Beginners' Intelligibility Test (BIT; Osberger et al., 1994), developed for assessing connected speech intelligibility in children with cochlear implants, replicate the results obtained with other measures with respect to full intelligibility in children with normal hearing by approximately age 4 years? That is, can convergent validity be established for the BIT with respect to other tests of intelligibility? Second, using the BIT, how intelligible are children who use cochlear implants in comparison to children of the same age with normal hearing?

The latter research question ignores for the time being the fact that children who use cochlear implants may have identical chronological ages but different lengths of experience with the implant. Previous research has determined that communication abilities increase with increased length of use of a cochlear implant, but children may be fitted with a cochlear implant at different chronological ages. Comparisons of children with like chronological ages may thus be confounded by the differences in length of experience with a cochlear implant. However, with increasingly higher expectations of benefits from cochlear implantation and increasing pressure toward educational mainstreaming and placement with same-age peers, it is important to know whether the communicative abilities of a child with a cochlear implant are age appropriate, that is, to know whether a child with a cochlear implant who looks and walks like an x-year-old also talks like an x-year-old.

To address this question, we administered the Beginners' Intelligibility Test (BIT; Osberger et al., 1994) to a group of children with normal hearing between the ages of 2 and 11 to determine approximate age-level performance on this assessment of connected speech intelligibility. These scores were then compared to archival data from the same task administered to an age-matched group of young children with profound deafness and cochlear implants. This comparison provides a context for the assessment of speech intelligibility in children who use cochlear implants that is different from contexts previously reported and has implications for expectations regarding outcomes of pediatric cochlear implantation, for educational placement and services, and for remediation.

Method

Participants

Participants were children with normal hearing, children with profound deafness and cochlear implants, and adults with normal hearing acting as listener judges of intelligibility.

Children with Normal Hearing

Connected speech intelligibility data were collected from a group of children who attended the Center for Young Children, a day-care facility on the campus of Indiana University–Purdue University Indianapolis. Children were recruited by letters to parents or guardians requesting their children's participation as members of a comparison group in a study examining the "development of speech" of children with hearing impairments. Speech data were collected from all children whose parents accepted the invitation to participate. However, based on responses on a demographic questionnaire, only data from children who met the following inclusionary criteria were analyzed: (a) no known speech or hearing problems, and (b) English as a native language. Data from children who lived in a home where a second language also was spoken were not excluded from analysis.

Using the foregoing inclusion criteria, data from 50 children were analyzed for Study 1 and from 34 children (a subset of the 50) for Study 2. For the group of 50 children in Study 1, age at time of testing ranged from 2;6 to 11;1 (*M* = 4;10, *SD* = 1;7; median = 4;7). For the group of 34 children for Study 2, age at time of testing ranged from 3;1 to 6;9 (*M* = 4;8, *SD* = 0;11; median = 4;7).

Children with Cochlear Implants

Connected speech intelligibility data from pediatric cochlear implant users were extracted from the database maintained in the DeVault Otologic Research Laboratory of the Department of Otolaryngology–Head and Neck Surgery at the Indiana University School of Medicine. These data had been collected prior to inception of the comparison study with normal-hearing children as part of a larger NIH-supported research project examining the speech and language development of children with cochlear implants.

Demographic information for each of the 34 children who used cochlear implants is included in the appendix. Children's age at onset of deafness ranged from 0;0 to 2;1 (*M* = 0;1, *SD* = 0;5), and their age at the time of receiving a cochlear implant ranged from 1;4 to 5;3 (*M* = 3;1, *SD* = 1;0). Unaided pure-tone average thresholds in the better ear before implantation ranged from 90 to 120.07 dB HL (*M* = 110.53, *SD* = 7.33; median = 111.69). Age at time of testing ranged from 3;1 to 6;9 (*M* = 4;8, *SD* = 0;11; median = 4;7), and length of cochlear implant use ranged from 0.50 years to 3.90 years (*M* = 1.63 years, *SD* = 0.88 years; median = 1.5 years).

All of the children used currently provided processing strategies. Three of the children used the advanced combination encoding (ACE) speech coding strategy with a Nucleus CI24M cochlear implant, four used the continuous interleaved sampling (CIS) strategy with a Clarion implant, and 27 used the spectral peak (SPEAK) strategy (25 with a Nucleus-22 device and 2 with a Nucleus CI24M); see Wilson

(2000) for a review of cochlear implant speech processing strategies. Nineteen of the children used oral-only communication, and 15 used total communication (a combination of spoken and signed language).

Listener Judges

Listener judges of the connected speech intelligibility of the children with cochlear implants and the children with normal hearing were adults between the ages of 18 and 40 who reported normal speech and hearing, as well as English as their native language. Potential judges of the speech of children with cochlear implants were excluded if they had more than minimal experience with speech produced by persons with hearing impairment. Listeners were recruited by means of two mail lists on the campus of Indiana University–Purdue University Indianapolis, one distributed to all students, faculty, and staff, and the other distributed specifically to medical students. Listeners were paid for their participation.

Materials

Connected speech intelligibility data were collected from all children participating in this study using the Beginners' Intelligibility Test (BIT; Osberger et al., 1994; see also Miyamoto, Kirk et al., 1997; Miyamoto, Svirsky et al., 1997). The BIT uses small objects and pictures to convey the context of the target sentence, and the child is instructed to produce an imitative response to the examiner's spoken model. Sentences in the test contain words that would be familiar to young children, simple syntactic structure is used in all sentences, and none of the words used in the sentences is more than two syllables long. Sentences range in length from 2 to 6 words (M = 3.8 words) and from 3 to 7 syllables (M = 4.5 syllables). Each of the four lists of 10 sentences contains between 37 and 40 total words (M = 38.3 words).

Procedures

Elicitation and Recording

Each child was administered one 10-sentence list from the BIT. Administration for children with cochlear implants was conducted in small testing rooms in the DeVault Otologic Laboratory at the Indiana University Medical Center in Indianapolis. Administration for children with normal hearing was conducted in small meeting rooms or empty classrooms at the Center for Young Children on the campus of Indiana University–Purdue University Indianapolis. All sessions with children were audio-recorded onto high-quality cassette tapes using lavalière microphones and a Marantz PMD430 recorder.

Listener-tape Preparation

Sentence productions from children who used cochlear implants were digitized using CSpeechSP (Milenkovic & Read, 1997) with a 22-kHz sampling rate, 16 bits per sample. The digitized sentences were then edited to remove extraneous material (e.g., examiners' models). Sentence productions from children with normal hearing were similarly digitized and edited using CoolEdit 2000 software. In addition, a small amount of background noise on a few tapes resulting from nonoptimal siting of testing rooms was digitally filtered and removed. For each child, a batch file was created consisting of two repetitions of each of the 10 sentences, along with stimulus cues ("Number X: ready" and "Number X again: ready"). ISIs were 2 seconds between repetitions of the same sentence and 4 seconds between different sentences.

Listener tapes were produced by recording the output of the digital batch files to high-quality cassette audiotapes using a Nakamichi BX-3000 discrete-head cassette deck. One tape was created for each panel of listeners, so each tape contained up to four BIT lists and no more than one recitation of each of the four BIT lists.

Listener Judgments

The speech intelligibility of each child was judged by three adult listeners with normal hearing. Stimuli were presented in a sound field at 65 to 70 dB HL to the listeners, who were seated in a sound-attenuated booth. Listeners were supplied with paper and pencil and instructed to write down everything that they heard the children say. Each panel of listeners heard either only children with cochlear implants or only children with normal hearing. Scores for individual children were calculated as the mean number of whole words understood correctly across the three listeners and converted to percent words correct.

Analysis

For Study 1, data from the group of 50 children with normal hearing were analyzed to determine the correlation between age at testing and BIT score. Additionally, differences in BIT scores between nominal age groups (e.g., age group 2 included children between 2;0 and 2;11) were analyzed by ANOVA. Study 2 was a comparison of children with normal hearing and an age-matched group of children who used cochlear implants. For each child with normal hearing, data from a child of the same chronological age were selected from the archival database of children who used cochlear implants. Selection was blind with respect to the recorded BIT score, and no child using a cochlear implant contributed more than one data point. Matching was possible in only 34 cases, so that Study 2 analyzed data from 34 children with normal hearing and 34 children with cochlear implants. The analyses in Study 2 were similar to those in Study 1, that is, correlational analyses of age and BIT score and comparison of mean BIT scores for different age groups using ANOVA. Additionally, mean BIT scores across like age groups for the two groups of children were compared.

Results

Study 1: Children with Normal Hearing

Scores on the BIT ranged from 13.5% to 100% correct (M = 87.2%, SD = 19%), with a median score of 94.7% correct. The scatterplot and regression line in Figure 1 show the distribution of BIT scores by ages of the children. For the children with normal hearing included in Figure 1, there was a moderate significant correlation between age at time of testing and BIT score ($r = .504$, $p < .00001$).

Descriptive statistics for BIT scores of nominal age groups (e.g., 2 = 2;0 to 2;11) of the 50 children with normal hearing are shown in Table 1. The data in Table 1 indicate that both mean and median scores on the BIT tended to increase as the age group increased. This was true for age groups 2 through 6, although the mean and median scores showed a slight decline from age group 6 to age group 7+.

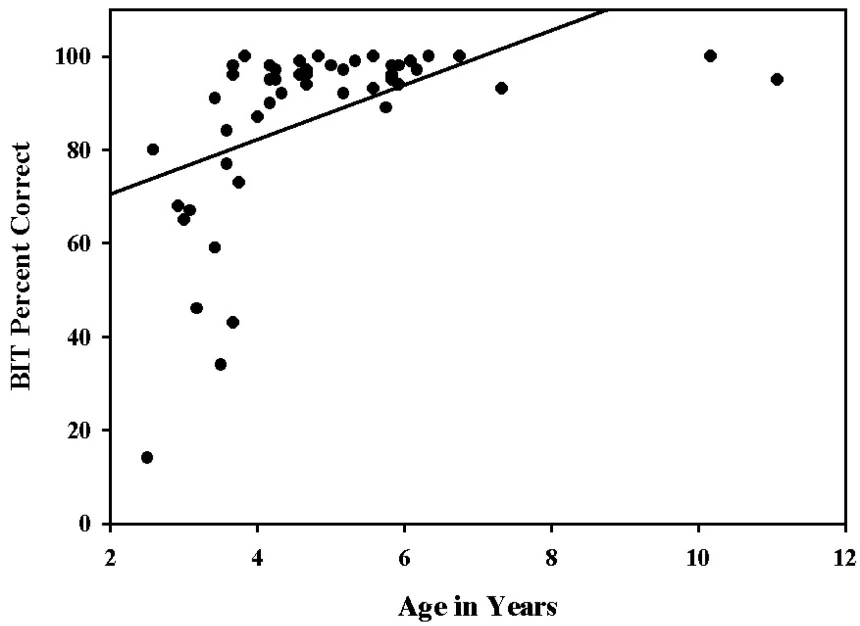


Figure 1. Scatterplot and regression line for ages in years (x-axis) and BIT percent correct scores for 50 children with normal hearing

Age Group (Median age)	N	BIT Score (percent correct)				
		Min	Max	Mean	SD	Median
2 (2;7)	3	13.5	79.8	53.9	35.5	68.4
3 (3;7)	13	34.2	100	71.8	21.9	72.8
4 (4;4)	13	86.8	100	95.2	3.75	96.5
5 (5;7)	14	89.2	100	96.2	3.15	97.0
6 (6;1)	4	97.4	100	99.1	1.24	99.6
7+ (10;2)	3	92.5	100	95.7	3.87	94.6

Table 1. Descriptive statistics for BIT scores of 50 children with normal hearing

To test for significant differences in BIT score means of the different age groups, a one-way analysis of variance with age group as the between-groups factor was conducted. The ANOVA revealed a significant effect of age, $F(5, 44) = 9.39$, $p < .001$. Post hoc multiple comparisons (Student-Newman-Keuls method) to determine which pairs of age groups were significantly different indicated that the mean score for age group 2 differed significantly from the means for age group 4 ($q = 6.514$, $p < .001$), age group 5 ($q = 6.726$, $p < .001$), age group 6 ($q = 5.987$, $p < .01$), and age group 7 ($q = 5.176$, $p < .01$). Similarly, the mean score for age group 3 differed significantly from the means for age group 4 ($q = 6.038$, $p < .001$), age group 5 ($q = 6.524$, $p < .001$), age group 6 ($q = 4.843$, $p < .05$), and age group 7 ($q = 3.782$, $p < .05$). Conversely, there was no significant difference between means for age groups 2 and 3, nor were there significant differences between pairs of means for age groups 4, 5, 6, and 7. These results thus indicated a threshold of speech intelligibility for children with normal hearing occurring between ages 3 and 4 years.

Study 2: Comparison of Children with Normal Hearing and Children with Cochlear Implants

Children with Normal Hearing

The group of 34 children with normal hearing was a subset of the larger group of 50. Comparison by *t* tests indicated no significant differences between the mean age of each age group in the 34 children and the corresponding age group in the larger group of 50. For the 34 children with normal hearing, scores on the BIT ranged from 43.3% to 100% correct ($M = 90.1\%$, $SD = 15.0\%$), with a median score of 96.4%. Descriptive statistics for the BIT scores of nominal age groups (e.g., 2 = 2;0 to 2;11) of the 34 children with normal hearing are shown in Table 2.

Age Group (Median age)	N	BIT Score (percent correct)				
		Min	Max	Mean	SD	Median
3 (3;8)	9	43.3	100	73.9	22.0	72.8
4 (4;4)	11	86.8	100	95.0	4.1	96.5
5 (5;7)	11	89.2	100	96.1	3.51	97.3
6 (6;2)	3	97.4	100	98.8	.3	99.1

Table 2. Descriptive statistics for BIT scores of 34 children with normal hearing

The data in Table 2 indicate that both mean and median scores on the BIT tended to increase as the age group increased. This is further illustrated by the scatterplot and regression line in Figure 2. For the children with normal hearing included in Figure 2, there was a modest but significant correlation between age at time of testing and BIT score ($r = .586$, $p < .001$).

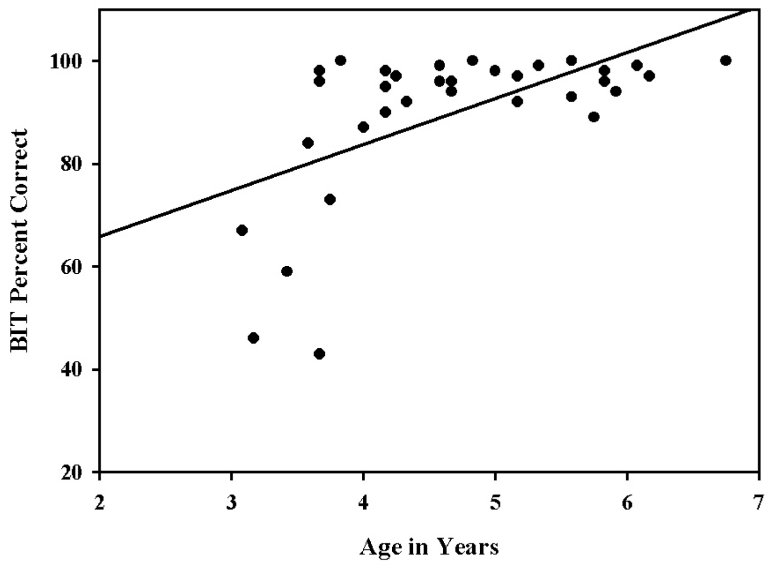


Figure 2. Scatterplot and regression line for ages in years (x-axis) and BIT percent correct scores for 34 children with normal hearing

To test for significant differences in BIT score means of the different age groups, a one-way analysis of variance with age group as the between-groups factor was conducted. As with the larger group of 50, there was a significant effect of age group on BIT scores for the smaller group of 34 children with normal hearing, $F(3, 30) = 7.83$, $p < .001$. Post hoc multiple comparisons (Student-Newman-Keuls method) indicated that the mean BIT score for age group 3 differed significantly from those of age group 4 ($q = 5.463$, $p < .001$), age group 5 ($q = 5.938$, $p < .001$), and age group 6 ($q = 4.492$, $p < .05$). On the other hand, there were no significant pairwise differences in BIT scores for age groups 4, 5, and 6.

Children with Cochlear Implants

Mean length of cochlear implant use for the nominal age groups (e.g., age group 3 = 3;0 to 3;11) were as follows: age group 3: 1.24 years (*SD* = 0.58); age group 4: 1.40 years (*SD* = 0.45 years); age group 5: 2.03 years (*SD* = 1.07 years); age group 6: 2.13 years (*SD* = 1.58).

For all 34 users of cochlear implants, scores on the BIT ranged from 0% to 85% correct (*M* = 15.3%, *SD* = 23.1%), with a median score of 7% correct. The scatterplot and regression line in Figure 3 show the distribution of BIT scores by ages of the children. For the children with cochlear implants included in Figure 3, there was no significant correlation between age at time of testing and BIT score ($r = .235$, $p = .182$ [n.s.]).

Descriptive statistics for the BIT scores of nominal age groups of the 34 children who used cochlear implants are shown in Table 3. Table 3 shows that mean BIT scores increased as age group increased from 3 through 6. Median scores, however, did not exhibit a similar monotonic increase. To test for significant differences in BIT score means of the different age groups, a one-way analysis of variance with age group as the between- groups factor was conducted. The ANOVA showed that for the group of 34 children with cochlear implants, age group was not a significant factor in differences between BIT scores.

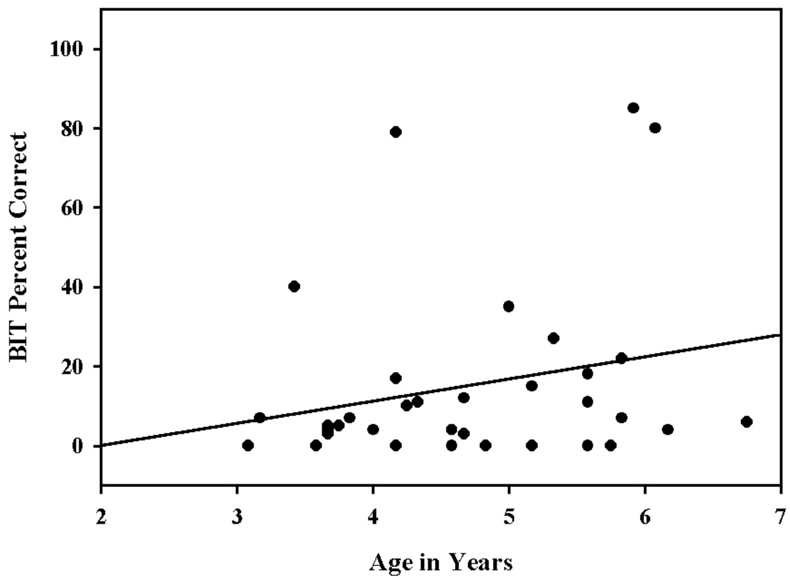


Figure 3. Scatterplot and regression line for ages in years (x-axis) and BIT percent correct scores for 34 children with cochlear implants

Age Group (Median age)	N	BIT Score (percent correct)				
		Min	Max	Mean	SD	Median
3 (3;8)	9	0.0	40.0	7.9	12.3	5.0
4 (4;4)	11	0.0	79.0	12.7	22.3	4.0
5 (5;7)	11	0.0	85.0	20.0	24.5	15.0
6 (6;2)	3	4.0	80.0	30.0	43.3	6.0

Table 3. Descriptive statistics for BIT scores of 34 children with cochlear implants

Comparison: Children with Normal Hearing and Children with Cochlear Implants

Table 4 shows mean BIT scores and standard deviations for the four age groups (3 to 6) for both children with normal hearing (left) and children with cochlear implants (right).

Age Group	Children with Normal Hearing		Children with Cochlear Implants	
	Mean	SD	Mean	SD
3	73.9	22.0	7.9	12.3
4	95.0	4.1	12.7	22.3
5	96.1	3.5	20.0	24.5
6	98.8	0.3	30.0	43.3
All	90.1	15.0	15.3	23.1

Table 4. Percent correct BIT scores for 34 children with normal hearing and 34 children with cochlear implants

To examine the effects of hearing status and age group on BIT scores, a 2 x 4 factoral ANOVA was conducted with hearing status (normal-hearing vs. cochlear implant) and age group (3 vs. 4 vs. 5 vs. 6) as between-groups factors. Results indicated main effects of both hearing status, $F(1, 60) = 203.53$, $p < .001$, and age group, $F(3, 60) = 3.94$, $p < .05$, but no interaction effect of hearing status and age group.

The plot in Figure 4 shows mean BIT scores by age group for both children with normal hearing (filled circles) and children with cochlear implants (open circles); the lines in this figure are included as a visual aid and not to imply linear increases in scores from circle to circle. As this figure shows, scores for children with normal hearing were consistently higher than scores for children with cochlear implants across all four age groups examined. The difference between percent correct scores for age group 3 was 66.0%; this increased at age group 4 (82.3%) but then decreased through age group 5 (76.1%) and age group 6 (68.8%).

Discussion

Consistent with previous studies (e.g., Coplan & Gleason, 1988; Gordon-Brannan & Hodson, 2000; Weiss, 1982; Weiss & Lillywhite, 1976), the children with normal hearing in Study 1 reached ceiling or near-ceiling around the age of 4 years. BIT scores for children 4;0 and older ranged from 86.8% to 100%, with a mean score of 96.1% (*SD* = .03%) and a median score of 97.0%. By contrast, scores for children with normal hearing below the age of 4;0 were considerably lower and more variable, ranging from 13.5% to 100%, with a mean score of 68.4% (*SD* = 24.6%) and a median score of 70.6%. Further evidence for an intelligibility threshold at age 4 is that there were significant differences in mean scores between ages 3 and 4 (and between ages 2 and 4) but no differences between age 4 and ages 5, 6, and 7+. In spite of the lack of significant differences in mean scores above age 4, there is nevertheless a significant correlation between age and BIT score across the range of ages tested. These generalizations also hold for the smaller group of 34 children with normal hearing analyzed in Study 2. With the group of 34 children who use cochlear implants from Study 2, the results indicate a somewhat different picture. There was no significant correlation between age and BIT score, and there were no significant differences in mean scores between any of the nominal age groups. Consequently, there was no obvious threshold of connected speech intelligibility among the children who used cochlear implants of the kind found for children with normal hearing.

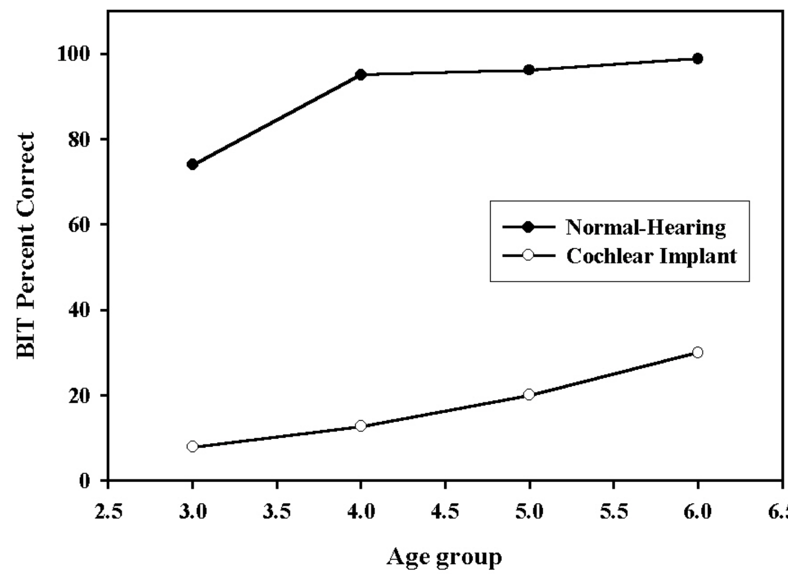


Figure 4. Mean percent correct BIT scores for age groups 3, 4, 5, and 6 for children with normal hearing and children with cochlear implants

The major difference between the children who use cochlear implants and the children with normal hearing, however, is that BIT scores for children with cochlear implants were significantly lower than for children with normal hearing. The 34 children with normal hearing had a mean BIT score of 90.1% (*SD* = 15.0%, median = 96.4%), whereas the age-matched group of 34 children with cochlear implants had a mean BIT score of 15.3% (*SD* = 23.1%, median = .07%). For the four nominal age groups, the difference in mean percent correct scores on the BIT ranged between 66.0% and 82.3%. The difference between the children with normal hearing and those with cochlear implants was lowest for age group 3 and highest for age group 4, indicating a large increase in scores for the children with normal hearing between ages 3 and 4 that was not present for the children with cochlear implants. From the large difference at age 4, the difference in scores then decreased through ages 5 and 6. This was due almost solely to increases in mean scores for the children with cochlear implants, the children with normal hearing having already achieved ceiling or near-ceiling scores.

The present study compares the connected speech intelligibility of children with normal hearing and children with cochlear implants solely on the basis of their chronological age. As the demographic data in the appendix show, the children who use cochlear implants examined here do not form a homogeneous group with regard to such characteristics as age at implantation, length of cochlear implant use, processing strategy, or communication mode. However, we believe that this very diversity serves the primary purpose of this study, namely to determine how different the speech intelligibility of children who use cochlear implants is from that of children with normal hearing of the same chronological age. As children with cochlear implants enter society and its institutions, a major factor affecting how well they will integrate is whether or not their speech intelligibility is consistent with other people's expectations based on experience with children of similar age with normal hearing. Using a different type of task, Monsen (1981) asserted that with intelligibility below 59%, "listeners are confronted with overwhelming difficulty in understanding what was said" (p. 850). In the present study, all but three children with cochlear implants had BIT scores below 59%, even up to the age of 6;9 (see Figure 3).

The results obtained in the present study, which specifically addresses the *age-appropriateness* of intelligibility before age 7, should not be construed to mean that children with cochlear implants cannot become intelligible. Figure 3 shows that some children with cochlear implants have relatively high speech intelligibility before age 7: 79% for SMY at age 4;2, 85% for SMC at age 5;11, and 80% for SKT at age 6;1. Furthermore, both SMC and SKT had durations of implant use that were considerably above the mean for this group (3.0 years and 3.9 years, respectively, vs. a mean of 1.63 years), and for this group of children, BIT scores were correlated significantly with length of device use ($r = .64$, $p < .0001$). Thus, although the majority of the children with cochlear implants in the present study had speech intelligibility scores well below the scores of same-aged children with normal hearing, there was still a tendency for scores to increase with age and a signficant correlation between scores and duration of use. Finally, the youngest age at implantation in this group was 1;4, but recent increases in the number of children given implants before age 1;0 will almost certainly result in higher speech intelligibility scores at ages 2 through 6 than those reported here.

Conclusions

This study allows us to draw a number of conclusions regarding the speech intelligibility of both children with normal hearing and children with cochlear implants. First, on the Beginners' Intelligibility Test, children with normal hearing younger than age 4 years are quite variable in their speech intelligibility, but at age 4 or shortly thereafter, they achieve adult-like or near-adult-like intelligibility. This is consistent with the few previous studies that have contributed normative data on the development of intelligibility. Second, children with cochlear implants between ages 3 and 6 years are *on average* considerably less intelligent than their chronological-age peers with normal hearing. Third, however, at least up to age 6, the speech intelligibility of children with cochlear implants does increase with age and increased duration of device use. It will be the task of future research to determine if and when children with cochlear implants are able to achieve the same level of connected speech intelligibility that children with normal hearing achieve around the age of 4 years.

References

Allen, C., Nikolopoulos, T.P., Dyar, D., & O'Donoghue, G.M. (2001). Reliability of a rating scale for measuring speech intelligibility after pediatric cochlear implantation. *Otology & Neurology, 22,* 631-633.

Allen, M.C., Nikolopoulos, T.P., & O'Donoghue, G.M. (1998). Speech intelligibility in children after cochlear implantation. *American Journal of Otology, 19,* 742-746.

Chapman, R.S., Seung, H.K., Schwartz, S.E., & Kay-Raining Bird, E. (1998). Language skills of children and adolescents with Down syndrome: II. Production deficits. *Journal of Speech, Language, and Hearing Research, 41,* 861-873.

Chin, S.B., Finnegan, K.R., & Chung, B.A. (2001). Relationships among types of speech intelligibility in pediatric users of cochlear implants. *Journal of Communication Disorders, 34,* 187-205.

Clarke, W.M., & Hoops, H.R. (1980). Predictive measures of speech proficiency in cerebral palsied speakers. *Journal of Communication Disorders, 13,* 385-394.

Coplan, J., & Gleason, J.R. (1988). Unclear speech: Recognition and significance of unintelligible speech in preschool children. *Pediatrics, 82,* 447-452.

Dawson, P.W., Blamey, P.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 and Hearing, 16,* 551-561.

Gordon-Brannan, M. (1994). Assessing intelligibility: Children's expressive phonologies. *Topics in Language Disorders, 14*(2), 17-25.

Gordon-Brannan M., & Hodson, B.W. (2000). Intelligibility/severity measurements of prekindergarten children's speech. *American Journal of Speech-Language Pathology, 9,* 141-150.

Jerger, S. (1987). Validation of the pediatric speech intelligibility test in children with central nervous system lesions. *Audiology, 26,* 298-311.

Kent, R.D., Miolo, G., & Bloedel, S. (1994). The intelligibility of children's speech: A review of evaluation procedures. *American Journal of Speech-Language Pathology, 3,* 81-95.

Kent, R.D., Weismer, G., Kent, J.F., & Rosenbek, J.C. (1989). Toward phonetic intelligibility testing of dysarthria. *Journal of Speech and Hearing Disorders, 54,* 482-499.

Keuning, K.H., Wieneke, G.H., & Dejonckere, P.H. (1999). The intrajudge reliability of the perceptual rating of cleft palate speech before and after pharyngeal flap surgery: The effect of judges and speech samples. *Cleft Palate-Craniofacial Journal, 36,* 328-333.

Koegel, R.L., Camarata, S., Koegel, L.K., Ben-Tall, A., & Smith, A.E. (1998). Increasing speech intelligibility in children with autism. *Journal of Autism & Developmental Disorders, 28,* 241-251.

Kumin, L. (1994). Intelligibility of speech in children with Down syndrome in natural settings: Parents' perspective. *Perceptual & Motor Skills, 78,* 307-313.

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

Milenkovic, P.H., & Read, C. (1997). *CSpeechSP user's manual*. Madison, WI: Author.

Miyamoto, R.T., Kirk, K.I., Robbins, A.M., Todd, S., & Riley, A. (1996). Speech perception and speech production skills of children with multichannel cochlear implants. *Acta Otolaryngologica (Stockholm), 116,* 240-243.

Miyamoto, R.T., Kirk, K.I., Robbins, A.M., Todd, S., Riley, A., & Pisoni, D.B. (1997). Speech perception and speech intelligibility in children with multichannel cochlear implants. *Advances in Otorhinolaryngology, 52,* 198-203.

Miyamoto, R.T., Svirsky, M.A., Kirk, K.I., Robbins, A.M., Todd, S., & Riley, A.I. (1997). Speech intelligibility of children with multichannel cochlear implants. *Annals of Otology, Rhinology, & Laryngology, 106,* 35-36.

Mondain M., Sillon, M., Vieu, A., Lanvin, M., Reuillard-Artieres, F., Tobey, E., & Uziel, A. (1997). Speech perception skills and speech production intelligibility in French children with prelingual deafness and cochlear implants. *Archives of Otolaryngology–Head and Neck Surgery, 123,* 181-184.

Monsen, R.B. (1981). A usable test for the speech intelligibility of deaf talkers. *American Annals of the Deaf, 126,* 845-852.

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

O'Donoghue, G.M., Nikolopoulos, T.P., Archbold, S.M., & Tait, M. (1999). Cochlear implants in young children: The relationship between speech perception and speech intelligibility. *Ear and Hearing, 20,* 419-425.

Osberger, M.J., Maso, M., & Sam, L.K. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. *Journal of Speech and Hearing Research, 36,* 186-203.

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

Osberger, M.J., Robbins, A.M., Todd, S.L., Riley, A.I., & Miyamoto, R.T. (1994). Speech production skills of children with multichannel cochlear implants. In I.J. Hochmair-Desoyer & E.S. Hochmair (Eds.), *Advances in cochlear implants* (pp. 503-508). Vienna: Manz.

Samar, V.J., & Metz, D.E. (1991). [Reply:] Scaling and transcription measures of intelligibility for populations with disordered speech: Where's the beef? *Journal of Speech and Hearing Research, 34,* 699-704.

Subtelney, J. (1977). Assessment of speech with implications for training. In F. Bess (Ed.), *Childhood deafness* (pp. 183-194). New York: Grune & Stratton.

Svirsky, M.A. (2000). Speech intelligibility of pediatric cochlear implant users and hearing aid users. In S. B. Waltzman & N. L. Cohen (Eds.), *Cochlear implants* (pp. 312-314). New York: Thieme.

Svirsky M.A., & Chin, S.B. (2000). Speech production. In S.B. Waltzman & N.L. Cohen (Eds.), *Cochlear implants* (pp. 293-309). New York: Thieme.

Svirsky, M.A., Chin, S.B., & Miyamoto, R.T. (in press). Speech intelligibility of profoundly deaf pediatric hearing aid users. *The Volta Review*.

Svirsky, M.A., Sloan, R.B., Caldwell, M., & Miyamoto, R.T. (2000). Speech intelligibility of prelingually deaf children with multichannel cochlear implants. *Annals of Otology, Rhinology & Laryngology, 109*(12, Pt. 2), 123-125.

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. *American Journal of Otology, 12(Suppl.),* 165-173.

Tobey, E.A., Geers, A.E., Douek, B.M., Perrin, J., Skellett, R., Brenner, C., & Toretta, G. (2000). Factors associated with speech intelligibility in children with cochlear implants. *Annals of Otology, Rhinology, & Laryngology, 109*(12, Pt. 2), 28-30.

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

Vieu, A., Mondain, M., Blanchard, K., Sillon, M., Reuillard-Artieres, F., Tobey, E., Uziel, A., & Piron, J. P. (1998). Influence of communication mode on speech intelligibility and syntactic structure of sentences in profoundly hearing impaired French children implanted between 5 and 9 years of age. *International Journal of Pediatric Otorhinolaryngology, 44,* 15-22.

Weiss, C.E. (1982). *Weiss Intelligibility Test*. Tigard, OR: CC Publications.

Weiss, C.E., & Lillywhite, H.S. (1976). *Communicative disorders*. St. Louis: C. V. Mosby Co.

Weston, A.D., & Shriberg, L.W. (1992). Contextual and linguistic correlates of intelligibility in children with developmental phonological disorders. *Journal of Speech and Hearing Research, 35,* 1316-1332.

Wilson, B.S. (2000). Strategies for representing speech information with cochlear implants. In J. Niparko, K.I. Kirk, N.K. Mellon, A.M. Robbins, D.L. Tucci, & B.S. Wilson (Eds.), *Cochlear implants: Principles and practices* (pp. 129-170). Philadelphia: Lippincott Williams & Wilkins.

Appendix

Demographic information for 34 pediatric users of cochlear implants

Code[a]	PTA[b]	Age at Onset of Deafness	Age at Implantation	Age at Test	Length of CI Use (years)	Strategy[c]	Mode[d]
SIV	111.70	0;0	2;7	3;1	0.5	CIS	TC
SLF	103.37	0;0	2;4	3;2	0.8	SPEAK	OC
SHJ	116.67	0;0	1;4	3;5	2.0	SPEAK	OC
SHV	116.73	0;0	3;0	3;7	0.6	ACE	TC
SKZ	118.40	2;1	2;9	3;8	0.9	SPEAK	OC
SMD	118.37	0;0	2;6	3;8	1.2	SPEAK	TC
SMH	118.43	0;0	2;2	3;8	1.5	SPEAK	OC
SLH	111.70	0;0	2;1	3;9	1.7	CIS	OC
SWJ	105.03	0;0	1;9	3;10	2.0	SPEAK	OC
SLD	103.33	0;0	3;3	4;0	0.7	CIS	OC
SWC	98.33	0;0	2;1	4;2	2.1	SPEAK	OC
SMT	108.33	0;0	3;2	4;2	1.0	SPEAK	TC
SMY	106.67	0;0	2;2	4;2	2.0	SPEAK	OC
SMR	111.70	0;0	2;9	4;3	1.5	SPEAK	OC
SNA	120.07	0;0	2;9	4;4	1.5	SPEAK	TC
SNS	118.43	0;4	3;0	4;7	1.6	ACE	TC
SMK	105.33	0;0	3;7	4;7	1.0	SPEAK	OC
SGD	111.67	0;4	3;9	4;8	0.9	SPEAK	OC
SMI	115.03	0;0	3;1	4;8	1.6	SPEAK	OC
SIO	115.07	0;0	3;3	4;10	1.5	SPEAK	TC
SIU	118.43	1;6	2;0	5;0	3.0	SPEAK	TC
SMO	110.00	0;0	4;8	5;2	0.5	SPEAK	TC
SMJ	110.07	0;0	3;7	5;2	1.6	SPEAK	TC
SMB	103.33	0;0	2;2	5;4	3.1	SPEAK	OC
SIY	118.43	0;0	2;7	5;7	3.0	SPEAK	TC
SND	110.00	0;0	3;6	5;7	2.1	SPEAK	OC
SHL	90.00	0;0	4;8	5;7	0.9	ACE	TC
SIC	118.40	0;0	5;3	5;9	0.5	SPEAK	TC
SLE	103.33	0;0	4;3	5;10	1.5	CIS	OC
SKN	103.33	0;0	2;9	5;10	3.0	SPEAK	OC
SMC	120.00	0;0	2;9	5;11	3.1	SPEAK	TC
SKT	106.67	0;0	2;2	6;1	3.9	SPEAK	OC
SKR	111.70	0;0	5;3	6;2	0.9	SPEAK	OC
SIN	100.00	0;0	5;2	6;9	1.6	SPEAK	TC

[a]Arbitrary triliteral code name (i.e., not participant initials); [b]Pure-tone average threshold in dB HL; [c]Speech processing strategy: SPEAK = spectral peak, CIS = continuous interleaved sampling, ACE = advanced combination encoder (see Wilson, 2000); [d]Communication mode: OC = oral-only communication, TC = total communication (combined spoken and signed communication)

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

Speech Perception Skills of Deaf Infants Following Cochlear Implantation: A First Report[1]

Derek M. Houston,[2] David B. Pisoni, Karen I. Kirk,[2] Elizabeth A. Ying[2] and Richard T. Miyamoto[2]

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

[1] Research supported by the American Hearing Research Foundation, IU Intercampus Research Grant, IU Strategic Directions Charter Initiative Fund, the David and Mary Jane Sursa Perception Laboratory Fund, and by NIH-NIDCD Training Grant DC00012 and NIH-NIDCD Research Grant DC00064 to Indiana University. We would like to thank Riann Mohar, Marissa Schumacher and Steve Fountain for assistance recruiting and testing infants; the faculty and staff at Wishard Memorial Hospital, Methodist Hospital, and Indiana University Hospital for their help in recruiting participants and the infants and their parents for their participation.
[2] Also DeVault Otologic Research Laboratory, Department of Otolaryngology-Head & Neck Surgery, Indiana University School of Medicine, Indianapolis, IN.

Speech Perception Skills of Deaf Infants Following Cochlear Implantation: A First Report

Abstract. We adapted a behavioral procedure that has been used extensively with normal-hearing (NH) infants, the Visual Habituation (VH) procedure, to assess deaf infants' discrimination and attention to speech. Twenty-four NH 6-month-olds, 24 NH 9-month-olds, and 16 deaf infants at various ages before and following cochlear implantation (CI) were tested in a sound booth on their caregiver's lap in front of a TV monitor. During the *habituation phase,* each infant was presented with a repeating speech sound (e.g., "hop hop hop") paired with a visual display of a checkerboard pattern on half of the trials ("sound trials") and only the visual display the other half ("silent trials"). When the infant's looking time decreased and reached a habituation criterion, a *test phase* began. This consisted of two trials: an "old trial" that was identical to the "sound trials" and a "novel trial" that consisted of a different repeating speech sound (e.g., "ahhh") paired with the same checkerboard pattern. During the habituation phase, NH infants looked significantly longer during the sound trials than during the silent trials. However, deaf infants who had received cochlear implants (CIs) displayed a much weaker preference for the sound trials. On the other hand, both NH infants and deaf infants with CIs attended significantly longer to the visual display during the novel trial than during the old trial, suggesting that they were able to discriminate the speech patterns. Before receiving CIs, deaf infants did not show any preferences. Taken together, the findings suggest that deaf infants who receive CIs are able to detect and discriminate some speech patterns. However, their overall attention to speech sounds may be less than NH infants. Attention to speech may impact other aspects of speech perception and spoken language development, such as segmenting words from fluent speech and learning novel words. Implications of the effects of early auditory deprivation and age at CI on speech perception and language development are discussed.

Introduction

Recent advances in cochlear implant (CI) technology have allowed an increasing number of deaf individuals to have access to sound and audition. For prelingually deaf children, CIs represent a novel sensory input, which provides a means for learning spoken language. The success of cochlear implantation in enabling deaf children to learn spoken language has led to a broadening of candidacy criteria to include younger and younger children. In 2000, FDA guidelines approved cochlear implantation for children as young as 1-year of age, and some surgeons are providing CIs to infants under 1-year of age when there is the availability of clear evidence that they are not receiving benefits from conventional hearing aids.

The population of early-implanted deaf infants is likely to increase substantially because of broadening of candidacy criteria and because hearing loss can now be detected at younger ages due to new screening methods. Position statements and guidelines from the Joint Committee on Infant Hearing (2000) and the American Academy of Pediatrics (1999) have persuaded most state lawmakers to implement Universal Newborn Hearing Screening (UNHS) that requires hospitals to test the hearing of all newborns. Currently, 38 states in the U.S. and several countries throughout Europe and the rest of the world have adopted or will soon adopt UNHS.

The trend to detect and identify hearing loss and to provide interventions at younger and younger ages is driven by the general belief that the earlier in development a child has access to sound and hearing, the better the chances he or she will acquire spoken language skills that are comparable to normal-hearing (NH) children. Investigations conducted at a number of CI research centers provide support for this view. Several investigators have shown that deaf children who receive CIs at younger ages tend to perform better on spoken language comprehension and production tasks than deaf children who receive CIs at older ages (Fryauf-Bertschy, Tyler, Kelsay, & Gantz, 1997; Tobey, Pancamo, Staller, Brimacombe, & Beiter, 1991; Waltzman & Cohen, 1998). For example, in one recent study, Kirk and her colleagues (Kirk, Miyamoto, Ying, Perdew, & Zuganelis, in press) reported results of a study that measured receptive and expressive language skills of children every 6 months up to 2 years following cochlear implantation. They found that the rate of improvement on the language measures was greater for children implanted before 2 years of age than for children implanted between 2 and 4 years of age. Given these and other similar findings, it is reasonable to expect that providing CIs to deaf infants at even earlier ages (i.e., before the first year) will yield even greater benefits for prelingually deaf children.

As newborn hearing screening is implemented in hospitals across the United States, many more young infants will be identified with a hearing loss and these children will become potential candidates for CIs. This new population of CI recipients presents challenges to both researchers and clinicians who are concerned with evaluating the benefits of CIs in children. While steady progress has been made in developing behavioral techniques to evaluate the speech and language of children aged 2 years or older, measuring these skills in infants who are too young to follow verbal instructions is extremely difficult. The only current methods of assessing speech and language outcomes in infants who use sensory aids rely exclusively on parental-report questionnaires (Hayes & Northern, 1996). At the present time, it is not known if providing CIs at even younger ages will lead to even greater benefits in speech and language outcomes. It is important that researchers and clinicians develop new behavioral methodologies to measure the perceptual and linguistic skills of these young infants before and after they receive their CIs, and track, longitudinally, how these abilities develop and change over time.

In order to understand what kinds of speech perception and language skills we might expect from deaf infants following cochlear implantation, it is useful to consider how these skills typically develop in NH infants. Over the last 30 years, developmental scientists have used several behavioral procedures to investigate the perceptual and linguistic capacities of NH infants (Jusczyk, 1997; Werker et al., 1998). We briefly review some of these skills and discuss how attention to speech might be important for acquiring these skills. Then we report preliminary data on deaf infants' attention to and discrimination of speech sounds. The results reported below represent the first effort to measure and describe some of the fundamental speech perception capacities of deaf infants following cochlear implantation.

Speech Perception Skills during the First Year of Life

Speech Discrimination. The speech perception capacities that infants exhibit during the first six months of life appear to be general rather than language specific (Jusczyk, 1997). Infants are born equipped to learn any of the world's languages. During the first half-year, NH infants are able to detect and discriminate fine-grained differences in speech sounds that differentiate words in any of the world's languages. Numerous investigations have shown that young infants are able to discriminate vowels (Kuhl, 1983; Polka & Werker, 1994; Trehub, 1973) and consonants that differ with respect to voicing (Eimas, Siqueland, Jusczyk, & Vigorito, 1971), place (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, & Mehler, 1988; Jusczyk, Copan, & Thompson, 1978; Moffitt, 1971), and manner (Hillenbrand, Minifie, & Edwards, 1979; Miller & Eimas, 1983) of articulation. Moreover, up to about 8 months of age, infants are able to detect and discriminate many phonetic contrasts that are not phonologically relevant in the ambient language but are relevant in other languages (Best, 1995; Polka & Werker, 1994; Trehub, 1976; Tsushima et al., 1994; Werker & Tees, 1984; see Jusczyk, 1997 for a review).

During the second half of the first year of life, the initial, language-general, speech perception capacities develop into language-specific speech perception skills (Best, 1995; Polka & Werker, 1994; (Werker & Lalonde, 1988; Werker & Tees, 1984). For example, Werker and Tees (1984) tested English-learning in 6- to 8-month-olds and 10- to 12-month-olds' ability to detect sound contrasts that were distinctive in Hindi but not English. They found that 6- to 8-month-olds but not 10- to 12-month-olds were able to discriminate these specific contrasts, suggesting that sometime during the second six months of life, NH infants become less sensitive to acoustic phonetic characteristics of speech that are not distinctive in their native language. This attenuation of perceptual sensitivity to nonnative speech contrasts reflects an important shift from language-general to language-specific speech perception skills based on early experience in the language-learning environment. Learning about the organization and properties of speech sounds and speech patterns in the ambient language helps infants discover how to segment continuous speech into words and provides the fundamental basis for learning words and acquiring the grammar of the target language they are exposed to by their caregivers (Jusczyk, 1997).

Segmentation of Words from Fluent Speech. In written language, words are separated by unambiguous spaces on a page. By contrast, words in spoken language are not reliably marked by pauses or acoustic cues that are the same across talkers and speaking rates. Speech is a continuous acoustic signal when spoken to adults and when spoken to children and infants. This has been confirmed by investigations of caregivers' speech to infants that have shown that caregivers tend to speak to infants using fluent speech rather than speaking each word in isolation (van de Weijer, 1998; Woodward & Aslin, 1990). While adults automatically use their knowledge of words in the "mental lexicon" to facilitate word recognition in fluent speech (Suomi, 1993), infants may not have any words at all in memory to rely on. In order to build a vocabulary of the language, infants must develop perceptual skills that allow them to recognize and extract the sound patterns of words from the context of fluent speech and organize them in some systematic fashion in long-term lexical memory.

Over the past 10 years, developmental scientists have investigated the problem of segmentation in normal-hearing infants and the role of various types of linguistic cues to segmentation such as: rhythmic (Echols, Crowhurst, & Childers, 1997; Jusczyk, Houston, & Newsome, 1999), statistical/distributional (Goodsitt, Morgan, & Kuhl, 1993; Morgan & Saffran, 1995; Saffran, Aslin, & Newport, 1996), coarticulatory (Johnson & Jusczyk, 2001), phonotactic (Mattys & Jusczyk, 2001; Mattys, Jusczyk, Luce, & Morgan, 1999), and allophonic (Jusczyk, Hohne, & Bauman, 1999). Some of these cues, such as statistical/distributional and coarticulatory information may be similar across languages, while others vary substantially from language to language. Infants learn language-specific segmentation cues by discovering the organization of sounds in the ambient language. Peter Jusczyk and his colleagues have focused on the time when infants begin segmenting words in fluent speech and when they use language-specific information to influence their segmentation strategies. In their seminal study, Jusczyk and Aslin (1995) tested 6- and 7.5-month-olds ability to recognize words in fluent speech. During a *familiarization phase*, infants were presented with repetitions of two words presented in isolation (*cup* and *dog* or *bike* and *feet*). During a *test phase*, the infants were presented with four passages, two of which contained the familiarized words and two of which contained the unfamiliar target words. The 7.5-month-olds, but not the 6-month-olds, attended significantly longer to the passages with the familiarized words. Jusczyk and Aslin interpreted these results as evidence that by 7.5 months of age, infants are able to segment and recognize familiar words in fluent speech even after only a brief period of exposure.

During the second half of the first year of life, infants develop greater sensitivities to language-specific attributes of speech that may facilitate speech segmentation. In one study, Jusczyk, Cutler, and

Redanz (1993) investigated English-learning infants' sensitivity to the rhythmic properties of English words. Approximately 90% of content words in English begin with a stressed (or "strong") syllable (Cutler & Carter, 1987). Jusczyk et al. (1993) tested English-learning infants' preferences for lists of bisyllabic words that follow the predominant strong/weak stress pattern of English (e.g., doctor, candle) versus lists of bisyllabic words that follow a weak/strong stress pattern (e.g., guitar, surprise). They found that 9-month-olds but not 6-month-olds attended significantly longer to lists of words that followed the predominant stress pattern of English words – strong/weak. In a subsequent study, Jusczyk, Houston, and Newsome (1999) discovered that 7.5-month-old English-learning infants were able to segment strong/weak words from fluent speech but not weak/strong (also see Echols, Crowhurst, & Childers, 1997). Taken together, both sets of findings suggest that English learning infants' sensitivity to the rhythmic properties of words in their language plays an important role in their ability to segment words from fluent speech.

In addition to the rhythmic properties of spoken language, infants also attend to other language-specific aspects of speech that are useful for word segmentation. For example, by 9 months of age, infants appear to be sensitive to the phonotactic properties of speech (Friederici & Wessels, 1993; Jusczyk, Luce, & Charles-Luce, 1994). Phonotactics refers to how the sound segments (i.e., phonemes of language) are sequenced and ordered in different contexts. For example, the sequence /mt/ occurs more often between words than within words in English. Attention to these sequential properties of speech patterns can further inform infants about the types of sounds or sequences of sounds that are more likely to occur within or between words, which will contribute to more mature and sophisticated speech segmentation skills. Indeed, in a recent study, Mattys and Jusczyk (2001) found that by 9 months of age, English-learning infants can use phonotactic information to locate word boundaries.

At a somewhat more fine-grained level, variants (or allophones) of the same phoneme can also serve as word boundary cues (Bolinger & Gerstman, 1957; Church, 1987). For example, in English, aspirated stop consonants (such as the /t^h/ in "top") mark word beginnings because they do not occur in other word positions (Church, 1987). Jusczyk, Hohne, and Bauman (1999) found that 10.5-month-old, but not 9-month-old, English-learning infants treat two-syllable sequence as "nitrates" or as "night rates," depending on the variant of /t/ they hear. The findings of Jusczyk et al. (1999-a) in combination with the results reported by Mattys and Jusczyk (2001) suggest that infants' sensitivity to language-specific properties of phonemes, their variants, and the constraints on orderings influences how they segment words from fluent speech.

Recently, Jusczyk (1997, 2002) has proposed that English-learning infants who have normal hearing may initially begin segmenting the speech stream using rhythmic information and simply assume that every stressed syllable is a word onset. This may be a good "first-pass" strategy for word segmentation; breaking the input signal into smaller, more manageable sound patterns allows infants to notice the internal organization of segments and other language-specific properties (e.g., phonotactic and allophonic properties) at different locations within words. As infants integrate multiple cues and learn to segment words from fluent speech, they begin the process of learning words and acquiring a grammar of the language. In order to do this, they must not only hear the speech in their immediate environment and surroundings, but they must also attend to speech patterns so they can encode its organization and structure and begin to recognize the repetition of similar patterns on different occasions. NH infants automatically attend to and learn about the organization of sounds and sound patterns in their native language naturally, without any formal or explicit training. However, the same might not be true for congenitally deaf infants who have received CIs after some period of deafness. Because the early part of their development occurred with little if any auditory input, the neural mechanisms involved in speech perception, attention and learning may be quite different from those of NH infants.

Consequences of Early Auditory Deprivation on Development of Speech Perception Skills

The absence of sound during the first few months of life may affect neurobiological development at several points along the peripheral auditory pathway as well as other higher-level cortical areas. In a recent paper, Shepherd and Hardie (2001) reviewed findings relating to changes in the auditory pathway caused by deafness. At the level of the cochlea, deafness leads to degeneration of spiral ganglion cells as well as a reduction in the efficiency, spontaneous activity and temporal resolution of auditory nerve fibers (Shepherd & Hardie, 2001). At the level of the central auditory pathway, bi-lateral hearing loss results in reduction of synaptic density in the inferior colliculus. It is possible that auditory deprivation may affect auditory acuity post-cochlear implantation. Leake and her colleagues investigated this possibility using cats that were deafened for different lengths of time and then given cochlear implants (Leake & Hradek, 1988; Rebscher, Snyder, & Leake, 2001). They found that spatial selectivity of electrode impulses in the inferior colliculus was affected by the length of deafness and the amount of degeneration of the spiral ganglion cells.

At higher cortical levels, numerous studies over the last 40 years have shown that early sensory experience plays a critical role in the organization of the sensory cortices (see Kujala, Alho, & Naatanen, 2000 for a review). When input from one sensory modality is unavailable, regions in the brain that normally subserve that modality appear to become more responsive to inputs from other sensory modalities (Rauschecker & Korte, 1993). Helen Neville and her colleagues have investigated neural reorganization in humans using a variety of neural imaging techniques. They have found that some regions of the auditory cortex that only respond to auditory information in normal-hearing individuals respond to several different types of visual information in deaf individuals who have learned sign language (Neville & Bruer, 2001).

Intercortical projections also appear to be affected by sensory deprivation as well. In deaf cats, Kral and colleagues recorded responses from different layers of the auditory cortex. They found reduced synaptic activity in the infragranular layers, which are output to the other cortical regions (Kral, Hartmann, Tillein, Held, & Klinke, 2000). Ponton and Eggermont (2001) collected auditory evoked potentials from deaf children who use cochlear implants and found responses that are consistent with immature superficial cortical layers, which are important for intracortical and interhemispheric communication. Taken together, these findings on neural development suggest that auditory deprivation may impair or attenuate the development of neural pathways connecting the auditory cortex to other cortical areas of the brain. Connections between the auditory cortex and other cortices, particularly the frontal and prefrontal cortices, are important for establishing higher-level attentional and cognitive neural networks linked to auditory processing. Thus, to understand the effects of auditory deprivation on language development in young prelingually deaf infants, it is important to investigate not only auditory acuity after CI but also processes involved in perception, attention, learning, and other cognitive skills that may be affected by the absence of sound during early development.

Assessing Speech Perception Skills of Deaf Infants after Cochlear Implantation

Infants' attention to and discrimination of speech sounds is crucial for further language acquisition. To assess these skills in infant CI users, we have constructed a new research laboratory within the ENT Clinic at the Indiana University School of Medicine to assess the speech perception and language skills of deaf infants before implantation and at regular intervals following cochlear implantation. One of the procedures we have adapted for this research program is the Visual Habituation (VH) procedure, which has been used extensively for the past three decades to assess the linguistic skills of normal-hearing (NH) infants (Best, McRoberts, & Sithole, 1988; Horowitz, 1975; Polka & Werker, 1994). Our goals in this initial research were to: (1) validate VH with this population of deaf and hard-of-hearing infants, and (2) use VH to track and assess infants' attention to speech and measure their speech discrimination skills before and after receiving a cochlear implant.

Measuring and tracking the perceptual and linguistic development of young prelingually deaf infants who receive CIs is important for both clinical and theoretical reasons. From a clinical perspective, it is essential that new behavioral techniques be developed to assess the benefits of implanting deaf infants with CIs at very young ages and measure changes in benefit and outcomes over time after implantation. From a theoretical perspective, this research provides a unique opportunity to compare language development of normally hearing infants to deaf infants who have been deprived of auditory input and then have their hearing restored at a later age via a CI. Do these deaf children follow the same developmental time course as normal-hearing infants, even though their early auditory experience was radically different? Also, how does the initial absence of auditory information affect an infants' subsequent ability to attend to and acquire spoken language? These are important fundamental questions that address neural development and behavior.

Experiment: Visual Habituation Procedure

The VH procedure has been used extensively over the years to assess NH infants' ability to discriminate speech contrasts (e.g., Best et al., 1988; Polka & Werker, 1994). In the standard implementation of VH, infants are first habituated to several trials of a repeating speech sound presented simultaneously with a visual display (e.g., a checkerboard pattern) during a habituation phase. The same stimuli are presented on each trial, and the infant's looking times to the visual display are measured. When the infant's looking time decreases and he or she reaches a pre-determined habituation criterion, a novel auditory stimulus is presented with the same visual display that was used during habituation. An increase in looking time to the visual display when the novel auditory stimulus is presented is taken as evidence that the infant was able to detect the difference in the speech stimuli and respond to the novelty of the new sound pattern.

We have modified the VH procedure to assess infants' attention to speech as well as their speech discrimination skills. During the initial habituation phase, half of the trials include an auditory stimulus ("sound trials"). On the other half of the trials, the infants are presented with only the visual display ("silent trials"). By comparing infants' looking times to the visual display on sound and silent trials, we can obtain an objective measure of their attention to speech. In this study, both NH infants (6- and 9-month-olds) and deaf infants before and following cochlear implantation were tested to assess their attention abilities and to measure their speech discrimination skills. We believe that these basic skills are clinically relevant and important for understanding deaf infants' potential for perceiving speech and learning spoken language. Another important goal of this investigation was to validate the VH procedure with a clinical population of infants whose speech perception and language-processing skills are completely unknown.

Method

Participants. To date, we have tested 16 prelingually deaf infants who were enrolled in the IU Medical School's cochlear implant program. The data from two infants were excluded due to very poor auditory detection as measured by visual reinforcement audiometry. Of the remaining 14 infants, eight of the infants were tested prior to cochlear implantation (mean age = 10.9 months, range: 5.8 – 20.7 months). Seven were tested at least once at approximately 1 month post-CI (mean age = 16.4 months, range: 8.7 – 24.6 months), eight were tested at least once at approximately 3 months post-CI (mean age = 18.4 months, range: 9.6 – 27.3 months) and eight were tested at approximately 6 months post-CI (mean age = 20.4 months, range: 13.9 – 29.9 months). One participant (CI01), who was the youngest cochlear implant recipient at IU School of Medicine, received a CI at 6 months of age. We have followed CI01 closely and will report his individual data collected across multiple testing sessions. Finally, for comparison, we have also tested 24 NH 6-month-olds and 24 NH 9-month-olds.

Apparatus. The testing was conducted in a custom-made double-walled IAC sound booth. As shown in Figure 1, infants sat on their caregiver's lap in front of a large 55" wide-aspect TV monitor, which was used to present all of the visual and auditory stimuli. The experimenter observed the infant via a hidden closed-circuit TV camera and coded how long and in which direction infants looked by pressing keys on a computer keyboard. The experiments were implemented on the computer using the Habit software package (Cohen, Atkinson, & Chaput, 2000).

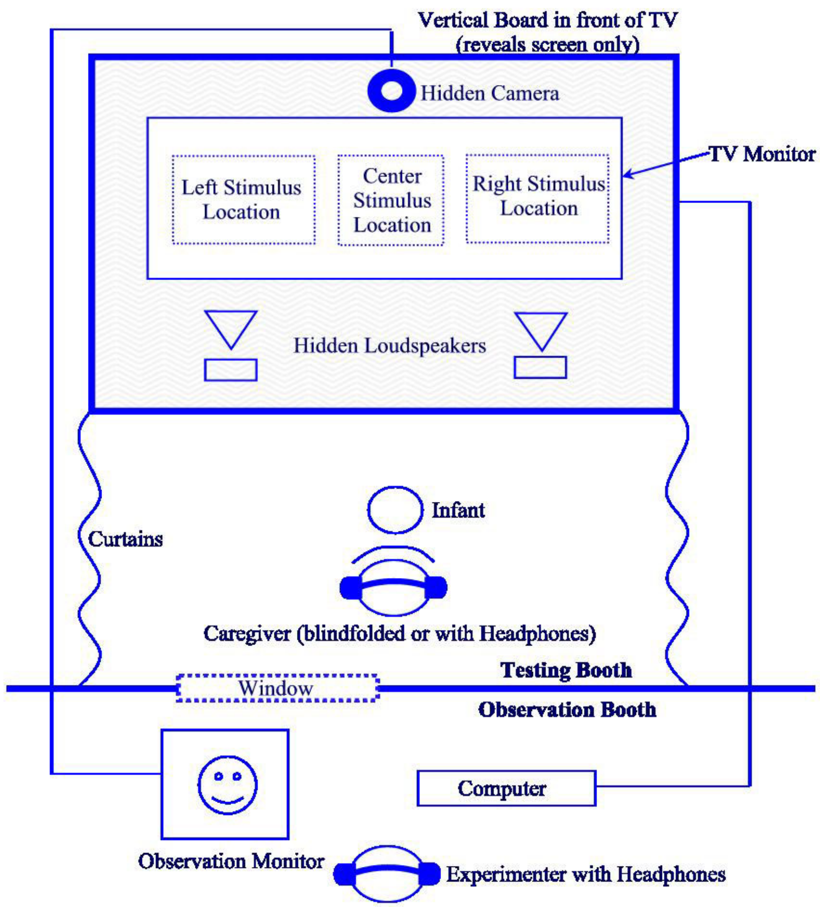


Figure 1. During the VH, the caregiver wears headphones playing masking music. The visual stimuli for the VH appear at the "Center Stimulus Location."

Stimulus Materials. To validate VH with this population of infants, we selected two very simple speech contrasts. These particular speech sounds are used clinically and have been found to be among the first sound contrasts that hearing-impaired children can detect and discriminate. One stimulus contrast was a 4 sec. continuous vowel ("ahhh") vs. 4 sec. discontinuous CVC pattern ("hop hop hop") contrast. The other contrast was a 4 sec. rising vowel /i/ vs. 4 sec. falling vowel /i/ intonation contrast. The stimuli were all produced by a female talker and recorded digitally into sound files that were presented to the infants at 70dB ± 5dB SPL via loudspeakers on the TV monitor. A computer representation of a red and white checkerboard pattern was created to serve as the visual display. Using VH, we assessed the ability to detect and discriminate these simple speech sounds in a group of congenitally deaf infants with CIs and a group of typically developing NH infants.

Procedure. The procedure we used was similar to the standard VH speech discrimination experiment. There was a habituation phase followed by a test phase. During the habituation phase, two types of trials were presented. *Sound trials* consisted of a pairing of the visual display and one of the sound stimuli (e.g., "hop hop hop" or "ahh"). *Silent trials* consisted of the visual display only with no sound presented. Two sound and two silent trials were presented, in random order, in each block of four trials. Infants' attention was initially drawn to the TV monitor was using an "attention getter" (i.e., a small dynamic video display of a laughing baby's face).

Each trial was initiated when the infant looked to the visual display. The trial continued until the infant looked away from the visual checkerboard display for 1 sec. or more. The duration of the infant's looking time toward the checkerboard was measured for each trial. During the habituation phase, the blocks of trials continued until the infant's average looking time to the visual display across a block of 4 trials (2 sound, 2 silent) was 50% or less than the average looking time across the first block of 4 trials. When this habituation criterion was met, the infant was then presented with two more trials, an old trial and a new trial (order of trials was counterbalanced across participants) during a test phase. The *old trial* was identical to the sound trials that the infant heard during the earlier habituation phase. The *novel trial* consisted of the other speech sound (e.g., "ahhh") of the pair and the same visual display. Based on previous research with NH infants, we predicted that if speech sounds elicited infants' attention then they would look longer to the visual display during the sound trials than during the silent trials. We also predicted that if the deaf infants could discriminate differences between speech sounds, they would exhibit longer looking times during the novel trial than during the old trial.

Results

We obtained two measures of performance. Attention was measured as the difference in the infants' looking times to the sound versus the silent trials. Speech discrimination was measured as the difference in the infants' looking times to novel versus old trials. Normal-hearing infants were only tested one time each, at either 6 months or 9 months of age. Deaf infants who received CIs were tested at several intervals before and after implantation. Data were grouped into 3 post-CI intervals: "One month" (1 day, 2 week, and 1 month post-CI intervals), "3 months" (2 and 3 month post-CI intervals), and "6 months (5 and 6 month post-CI intervals). Some of the deaf infants were tested more than once during a single interval group. In the final data analyses, each session was treated as an independent sample, rather than averaging across sessions within a post-CI interval group.

Attention to Speech Sounds. The looking times to the sound and silent trials were averaged separately for each infant. The average looking times were then subjected to a 3-way repeated-measures ANOVA with Auditory Condition (sound vs. silence) as a within-subjects factor and Stimulus Condition ("ahh", "hop hop hop", rising /i/, and falling /i/) and Group (Pre-implantation, post-CI, and NH) as between-subjects factors. There was no main effect of Stimulus Condition ($F < 1$) and no interactions of Stimulus Condition with any of the other factors (all $Fs < 1$), suggesting that the infants looking times were similar across the different stimulus conditions in the experiment. Based on these findings, we combined the data across the four stimulus conditions and re-analyzed the data using a 2-way repeated-measures ANOVA. Figure 2 displays the average difference in looking times (and 95% confidence intervals) to the sound versus silent trials for the NH 6- and 9-month olds (solid bars) shown on the left, the deaf infants pre-CI shown in the middle (striped bar), and the deaf infants at the 1-, 3-, and 6-month post-CI intervals (patterned bars) shown on the right. Bars above the line at zero represent longer looking times to the sound trials than the silent trials.

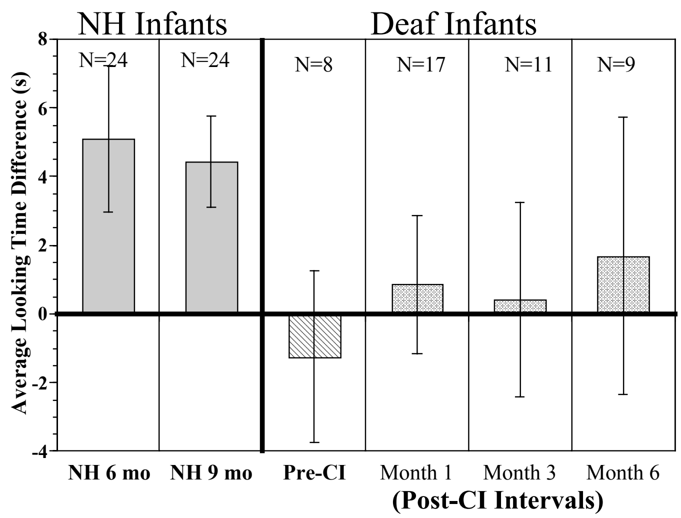


Figure 2. Attention to speech sounds. Looking time difference, averaged across stimulus conditions, to the sound versus the silent trials for normal-hearing (NH) controls, for deaf infants before cochlear implantation, and for deaf infants at several intervals after cochlear implantation (CI). The number of observations is given for each interval. Some deaf infants after implantation were tested more than once, yielding more observations than number of participants.

Overall, infants looked longer during the sound trials than during the silent trials (F (1, 90) = 6.79, p < .05). There was no main effect of group (F (5, 90) < 1), but there was a significant Group X Auditory Condition interaction (F (2, 90) = 13.21, p < .001), indicating, that not all of the groups looked significantly longer to the sound than to the silent trials. NH infants attended longer to the sound than to the silent trials (F (1, 46) = 62.51, p < .001), and there was no interaction with Group (6 and 9 months) (F (1, 46) < 1). Before receiving a CI, the deaf infants did not look longer to the sound trials than to the silent trials (t (7) = -1.17, p > .2). However, following cochlear implantation the deaf infants did attend longer to the sound than to the silent trials, although the overall looking time difference did not reach statistical significance (F (1, 34) = 1.79, p > .1). Also, the interaction between Auditory Condition and Deaf Group (pre-CI vs. post-CI) did not reach significance (F (1, 43) = 1.89, p > .1). Among the deaf infants post-CI, there was also no interaction between Auditory Condition and Post-CI Interval (month 1, 3, and 6) (F (2, 34) < 1), suggesting that the trend to look longer to the sound versus the silent trials was similar in magnitude across the post-CI intervals. Comparing the looking times of all NH infants with deaf infants post-CI revealed a significant Group X Auditory Condition interaction (F (1, 83) = 17.55, p < .001). This finding indicates that the difference in looking times between sound and silent trials was significantly greater for the NH infants than for the deaf infants who received CIs.

Figure 3 displays the data from deaf infant CI01. He was tested three different times between 1 and 3 months after receiving his cochlear implant. Over this time period, he showed very little difference in his looking times for sound versus silent trials. However, he was also tested five times between 6 and

15 months after cochlear implantation. Over this period, he displayed a consistent trend to look longer during the sound than silent trials (t (4) = 2.30, p = .08).

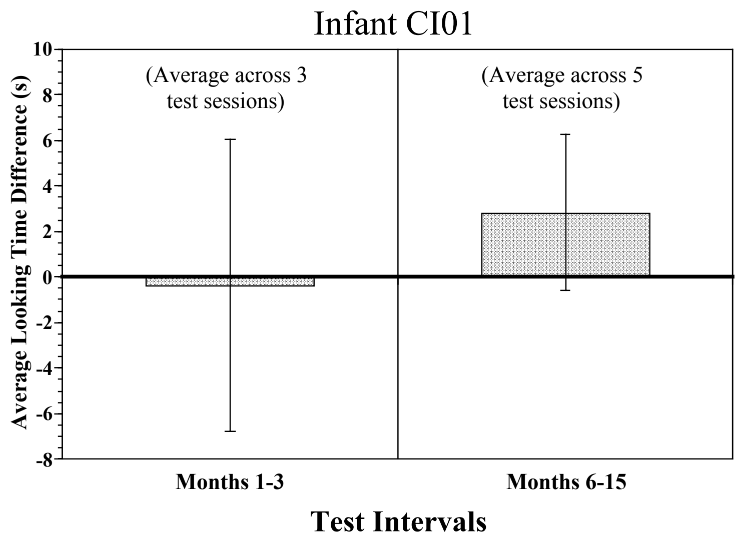


Figure 3. CI01 Attention to speech sounds. Looking time differences to the sound versus the silent trials for participant CI-01.

In summary, NH infants exhibited a strong preference for sound trials over silent trials. Before CI, deaf infants showed no such preference. After CI, deaf infants attended longer to the sound trials than the silent trials but much less so than the NH infants. In contrast, the deaf infant (CI01) who was implanted at 6 months of age, showed a preference for sound trials at his later CI intervals that was similar in magnitude to the NH infants' preference.

Speech Discrimination. The looking times to the novel and old trials were subjected to a 3-way repeated-measures ANOVA with Discrimination (novel vs. old) as a within-subjects factor and Stimulus Condition ("ahh" vs. "hop hop hop" and rising vs. falling /i/) and Group (Pre-CI, post-CI, and NH) as between-subjects factors. Figure 4 displays differences in looking times to the novel versus old trials for the NH infants shown on the left, the deaf infants pre-CI shown in the middle, and the deaf infants following CI shown on the right. The looking times are further divided by stimulus condition with the looking time differences for the "hop hop hop" versus "ahh" conditions indicated by the left bar in each panel and the rising versus falling /i/ indicated by the right bar within each panel in the figure.

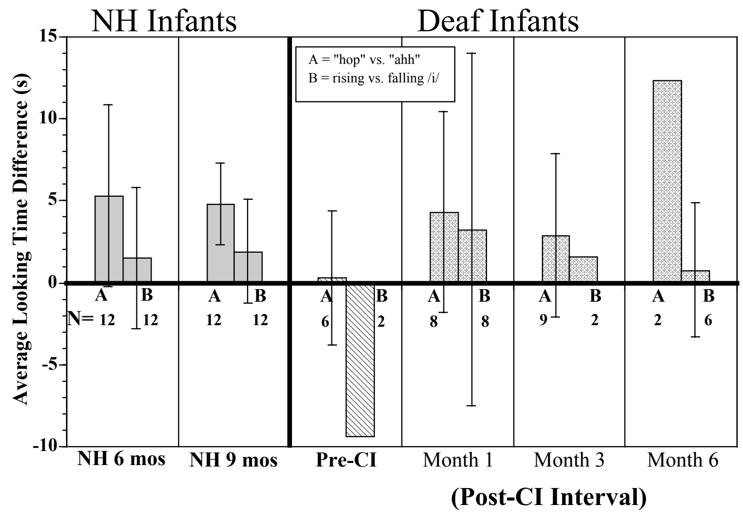


Figure 4. Speech discrimination. Looking time difference to the novel versus the old trial for normal-hearing (NH) controls and for deaf infants before and at several intervals after CI. Data are displayed separately by stimulus conditions: “hop hop hop” vs. “ahh” (A) and rising vs. falling /i/ (B). The number of observations is given for each interval. 95% confidence interval bars are displayed in cases where there were three or more observations.

Across groups, infants did not look longer to the new trials than the old trials (F (1, 85) < 1). However, there was a significant interaction between Group and Discrimination condition (F (2, 85) = 3.45, p < .05), suggesting that one of the groups may have discriminated the contrasts. There was also a significant interaction between Stimulus Condition and Discrimination condition (F (1, 85) = 5.49, p < .05), suggesting that the two pairs of stimuli were not equally discriminable. Further analyses revealed that deaf infants before CI looked longer to the old than to the new trials, but this difference did not reach significance (F (1, 6) = 3.22, p > .1). Analyses of the looking times of deaf infants at the three intervals after CI (month 1, month 3, and month 6) showed significantly longer looking times to the novel than to the old trials (F (1, 29) = 5.85, p < .05) and no significant interactions with Post-CI Interval (F (1, 29) < 1) or with Stimulus Condition (F (1, 29) = 1.80, p > .1). Likewise, NH infants attended significantly longer to the new trials than to the old trials (F (1, 44) = 13.57, p < .001), and there was no significant interaction with Age Group (F (1, 44) < 1). However, the interaction between Discrimination and Stimulus Condition approached statistical significance (F (1, 44) = 3.31, p < .08), reflecting a larger difference in discrimination for the “hop hop hop” versus “ahh” condition than the rising versus falling /i/ condition. Additional analyses comparing groups of infants revealed that the effect of discrimination was significantly different for the pre-CI and post-CI groups (F (1, 39) = 4.98, p < .05) and between the pre-CI and NH groups (F (1, 52) = 8.63, p < .01), but the interaction between post-CI and NH groups did not approach significance (F (1, 79) < 1). These results suggest that NH infants and deaf infants who received CIs discriminated both sound contrasts to similar degrees while the deaf infants pre-CI were unable to discriminate any differences reliably.

Figure 5 displays the looking times of deaf infant CI01 during early (1-3 month) and later (6-15 month) post-implantation intervals. Infant CI01 showed no preference for the novel stimulus during the early test intervals, but he did display a trend to look longer during the novel trials at later post-implantation intervals (t(4)=2.34, p=.08).

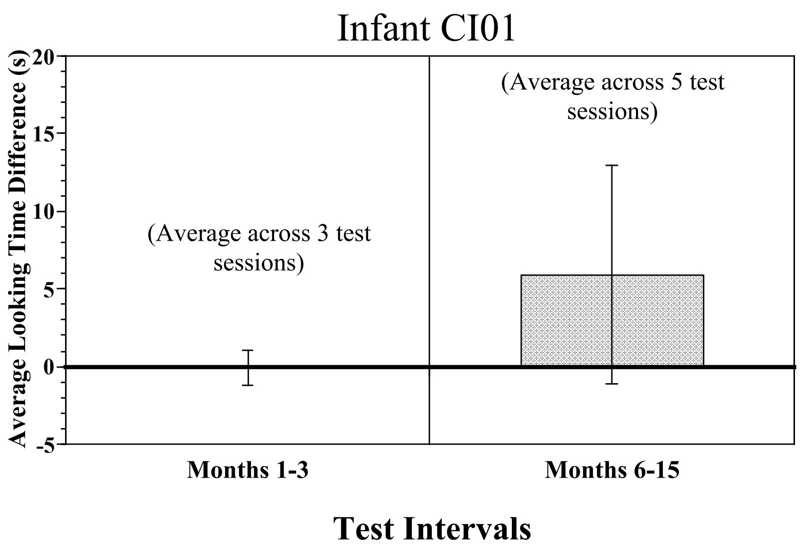


Figure 5. CI01 speech discrimination. Looking time differences to the novel versus the old stimulus trial for participant CI-01.

Discussion

The attrition rates observed in the VH task were similar across both groups of deaf and NH infants – about 20 to 25 percent. These rates are low-to-average compared to other speech perception experiments with NH infants (Werker et al., 1998). Hence, it appears that VH procedure is a viable behavioral technique that can be used with deaf infants before and after cochlear implantation to assess benefit and measure change in auditory attention and speech discrimination skills over time.

In the present investigation, deaf and NH infants' attention to speech sounds was assessed during a *habituation phase.* One of four repeating speech sounds was paired with a checkerboard pattern on half of the trials while the other half of the trials consisted of the checkerboard pattern with no sound. We observed no significant effect of stimulus type for any of the groups of infants, suggesting that when compared to silence, the infants were similarly interested in each type of speech stimulus. Preference for the sound trials over the silent trials differed across groups, however. Both 6- and 9-month-old NH looked significantly longer to the checkerboard pattern when accompanied by a repeating speech sound. Before implantation, deaf infants did not look longer during the sound trials. In contrast, after implantation, deaf infants did look longer during the sound trials, although the difference in looking times between sound and silent trials did not reach statistical significance.[3]

The NH infants' preference for the sound trials was significantly greater than the preference of the deaf infants with CIs even though the deaf infants' auditory detection thresholds (measured using visual reinforcement audiometry) were far below the intensity level of the stimuli (70 dB ± 5dB). In other words, although both groups of infants can detect the repeating speech stimuli, the NH infants increased their attention (i.e., looking times) much more in the presence of speech sounds than the deaf infants did following CI. These findings suggest that while deaf infants' attention to speech may increase slightly

[3] We expect that with a larger sample size, the difference in looking times between sound and silent trials for the deaf infants who use cochlear implants would reach statistical significance. However, their preference for sound trials would still be significantly smaller than the NH infants if the pattern of looking times was consistent with what has been observed so far.

during the first 6 months following CI, it does not reach the levels observed in NH infants. However, deaf infant CI01's preference for the sound trials at later intervals after CI was similar to the NH infants', suggesting that age at implantation and duration of CI use may contribute to the development of attention to speech. Infant CI01 was our youngest CI recipient. He received his CI at 6 months of age.

During the *test phase* of the VH experiment, we assessed infants' discrimination of two gross-level speech contrasts. NH infants looked significantly longer to novel sound trials than the old sound trials, validating this version of the VH procedure for use as a way of measuring speech discrimination. Similarly, deaf infants who received CIs demonstrated evidence of speech discrimination as well. Before receiving a CI, deaf infants did not look longer during novel sound trials than old sound trials. However, after receiving CIs, deaf infants did show a difference, suggesting that deaf infants' discrimination of some speech patterns improves with access to sound within the first few months after CI. For both the NH and deaf infants with CIs, the continuous ("ahh") and discontinuous ("hop hop hop") speech contrast appeared to be much more salient than the rising versus falling intonation pattern of /i/. Further investigations will examine if deaf infants who have received CIs also demonstrate discrimination when the acoustic differences in speech sounds are smaller, such as minimal pairs of words, which will provide more detailed information about the acuity of their speech perception skills for discriminating fine phonetic details in speech and spoken words.

Mean Looking Times

The difference scores used to assess looking times to the sound and silent trials plotted in Figure 2 provide a simple way to see the stimulus preferences of infants in the VH procedure. However, difference scores do not reveal how much time the infants actually looked at the visual displays for the two types of trials. To gain another perspective on the results, Figure 6 displays average looking times (and standard errors), rather than difference scores. In each panel, the bar shown on the left displays the average looking time to the sound trials, while the bar shown on the right displays the average looking time to the silent trials. This figure reveals that while the NH infants' and deaf infants' overall looking times were generally quite similar, the pattern of looking times differed in several important ways. First, NH infants looked longer to the sound trials than the deaf infants did. In contrast, NH infants looked less to the silent trials than the deaf infants did. These two results suggest that the overall attention level of the NH infants and deaf infants during the habituation trials was similar but that the NH infants attended more during the sound trials and less during the silent trials than the deaf infants who use CIs. During the test phase, when both trials were "sound" trials, NH infants' exhibited a trend to look longer during both the novel and old trials than the deaf infants, although this trend did not reach statistical significance. The infants' mean looking times during the test phase are displayed in Figure 7. Taken together, the pattern of findings is consistent with the hypothesis that the NH infants exhibit a greater degree of selective attention to sound than do the deaf infants following CI. Importantly, the results do not appear to be due to differences between the groups in general arousal or attention; the results are more selective in nature and are based on stimulus differences.

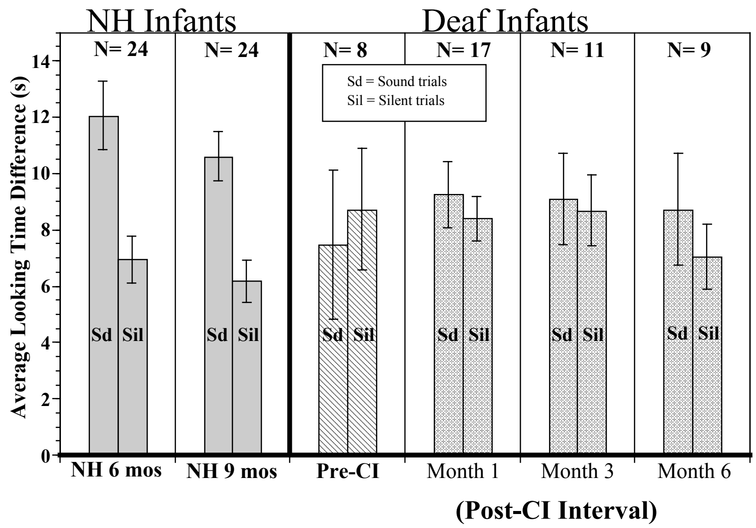


Figure 6. Attention to speech sounds. Mean looking times to the sound versus the silent trials for normal-hearing (NH) controls, for deaf infants before cochlear implantation, and for deaf infants at several intervals after CI.

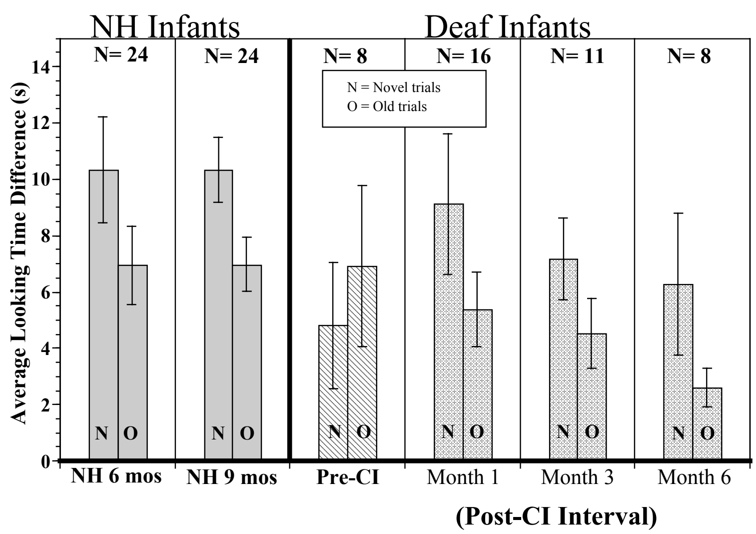


Figure 7. Speech discrimination. Mean looking times to the novel versus the old trial for NH controls and for deaf infants before and at several intervals after CI. Looking times are combined across stimulus conditions.

Attention to Speech

Comparing two different speech perception skills in deaf infants before and after CI with NH infants provides valuable new information about the immediate benefits deaf infants will likely gain from their CIs. These comparisons may also help researchers and clinicians understand the nature of the challenges these infants will face when acquiring spoken language. The results from this initial investigation provide some clues about deaf infants' speech perception skills after receiving their CI. The looking time response to gross-level changes in speech sounds after CI was similar to NH infants but was different from deaf infants pre-CI, demonstrating the development of fundamental attention and speech discrimination skills. In contrast, the pattern of looking behavior to sound versus silent trials in the deaf infants after CI was significantly different from NH infants. These findings suggest that even after receiving a CI, deaf infants do not appear to be as interested in speech as NH infants were in the VH procedure. However, several factors may be responsible for the findings that deaf infants post-CI did not respond similarly to NH infants in the attention task.

One factor that may have contributed to the differences in looking times between NH and deaf infants with CIs was the difference in chronological age between the two groups of infants. All of the NH infants used in this study were younger than the deaf infants at their post-CI intervals, except for CI01 at his month 1 post-CI interval. Thus, the NH infants were more closely matched to the deaf infants based on "hearing age" than on chronological age. This allowed us to compare, for example, NH 6-month-olds to older, deaf infants who have had 6 months of auditory experience with a CI. However, older infants may not show the same kinds of preferences for sound trials as young infants, regardless of hearing status. Hence, we cannot be sure if the differences in looking times we observed between deaf infants and NH infants were due to differences in early auditory experience or due simply to age differences. Further comparisons with older NH infants would provide more information about the differences between deaf infants and age-matched NH peers. However, the data collected so far from the deaf infants with CIs suggest that attention to speech will not diminish with age. In fact, infants who have used their CIs for longer periods of time and who were, on average, somewhat older than infants tested at earlier intervals showed increased preference for sound trials. Also, CI01's preference for sound trials consistently increased with age and CI experience.

Another factor that may have played a role in the pattern of results is the deaf infants' auditory acuity following CI. While CIs provide deaf individuals with access to sound, they do not provide nearly the richness of information that a healthy cochlea does. As a result, deaf infants who receive CI may not be able to detect and discriminate as many fine-grained details in speech as NH infants. As a consequence, the impoverished speech signal provided by a CI may be inherently less interesting than the speech signals processed by a healthy cochlea because deaf infants may not be able to detect, discriminate, and encode fine acoustic-phonetic features that represent the linguistically significant sound contrasts in the target language. Unfortunately, at the present time, very little is known about the auditory sensitivity and acuity of deaf infants in discriminating speech sounds. Further studies of speech discrimination by deaf infants with CIs will provide more knowledge of their speech perception skills and how discriminability and distinctiveness contribute to attention to speech sounds.

Developmental factors may also contribute to differences observed between the groups of infants. The deaf infants studied here developed with little, if any, exposure to sound until they received their CIs. Early sensory deprivation and lack of auditory experience with meaningful sounds as well as speech may have a significant impact on how infants learn to interact with objects and sound sources in their environment (Gaver, 1993). Neurophysiological studies provide evidence that sensory cortices are re-organized by early sensory experience (Neville & Bruer, 2001; Rauschecker & Korte, 1993) and that intercortical projections from the auditory cortex are affected by auditory deprivation (Kral et al., 2000; Ponton & Eggermont, 2001).

At this time, little is actually known about the extent of neural reorganization in human infants due to early auditory deprivation. And nothing is known about how absence of sound during early development might affect attention to speech after a child receives a CI. However, given that infants' ability to orient their attention to sensory input develops during the first year of life (Posner & Rothbart, 2000), it is quite possible that neural connections linking sensory perception to attention and other aspects of perception and cognition may not develop normally after a period of auditory deprivation and lack of stimulation. Moreover, before intervention deaf infants are not likely to respond to sound with activities such as visual orientation and vocal imitation. This lack of active response to sound may also affect the development of neural connections between auditory and other cortices.

Infants' attention to speech may have consequences for acquiring other speech perception skills that are important for learning spoken language. Perception and attention to fine-grained acoustic-phonetic details in speech are known to be important for distinguishing spoken words. Paying attention to the ordering of sounds in speech may play a role in learning about the organization of sound patterns in the native language. Deaf infants who do not maintain the same level of attention to speech that NH infants do may not develop normal sensitivities to language-specific properties, such as rhythmic, distributional, coarticulatory, phonotactic, and allophonic cues, all of which have been shown to be important sources of information for segmenting words from fluent speech and acquiring the vocabulary of a given language (Jusczyk, 1997; 2002). And, because infant-directed speech is typically continuous in nature (van de Weijer, 1998; Woodward & Aslin, 1990), difficulties segmenting words from fluent speech may cascade and produce atypical word learning and lexical development as well as morphological irregularities that affect syntax and language comprehension processes as well (Houston, Carter, Pisoni, Kirk, Ying, in press; Svirsky, Stallings, Lento, Ying, Leonard, 2002).

Some Future Directions

The present investigation, using the VH procedure, provides some preliminary data suggesting that deaf infants who have received CIs are able to discriminate gross-level speech sound contrasts but they appear to pay less attention to speech than NH 6- and 9-month-olds. One deaf infant (CI01), who was tested repeatedly over time, however, did pay more attention to sound trials than silent trials. It is possible that deaf infants who receive CIs at very young ages, like CI01 who received his CI at 6 months of age, will also demonstrate attention to speech that is more similar to NH infants. We hope to be able to more thoroughly assess the effects of age at implantation on attention and speech discrimination skills when we have collected more data from additional infants who have received CIs at different ages.

The sound contrasts that we studied in this investigation were gross-level, simple discriminations used to validate the VH procedure. After having established the VH as a viable measure of speech discrimination, it will be important to test deaf infants with acoustic-phonetic contrasts that are used to distinguish words in their native language. Measures of infants' perceptual sensitivity and acuity of speech in addition to other speech perception measures will be helpful in understanding the relation between the quality of their perception and how that cascades to affect other speech perception skills.

Several speech perception skills are important for segmenting words from fluent speech. Many of these skills involve acquiring sensitivity to language-specific statistical properties that are informative about the organization of sounds in the native language. Assessing deaf infants' sensitivity to these sequential properties will provide valuable new information about the kinds of information they attend to in speech. If attention to speech underlies the sensitivity to language-specific properties, then we might expect infants who show poor attention to speech to also display poor sensitivity to language-specific properties relative to NH peers, even when the NH infants are matched for the amount of time they have had access to sound. Moreover, sensitivity to language-specific properties may be important for speech segmentation and later word learning. Assessing these skills in deaf infants at several intervals after implantation and comparing individual infants' skills using different speech perception and word-learning measures may inform us about the links between attention and speech, sensitivity to language-specific properties, word segmentation, and word learning and the effects of auditory deprivation on the development of these skills.

Finally, while it is important to acquire detailed knowledge about the development of the receptive skills necessary for segmenting and identifying words in fluent speech, these skills alone are not sufficient for learning language. Children must also develop the expressive productive skills necessary for forming and articulating intelligible utterances. The linguistic environment plays an important role in the development of NH infants' early productive skills. Investigations have shown that the particular language NH infants are exposed to influences the segmental (Boysson-Bardies et al., 1992) and the rhythmic (Levitt, Utman, & Aydelott, 1992) characteristics of their babbling by the end of the first year of life. The language learner's ability to produce language is also influenced by input from the visual modality. Kuhl and Meltzoff (1982) reported that 20-week-old infants attended longer to a video display of a speaker articulating the vowel they were hearing than to a video of the same speaker articulating a different vowel, suggesting that NH infants are able to detect and use cross-modal correspondences between the auditory and visual properties of speech. Noticing auditory and visual correspondences may be integral to NH infants' imitations of sounds in their environment. Their own vocal imitations may, in turn, provide infants with auditory feedback that they can use to adjust their productions to more closely resemble utterances in the target language.

For profoundly deaf infants, the opportunity to integrate auditory and visual information is absent until auditory information is made available by CI. Consequently, their ability to integrate auditory and visual information may be delayed or impaired, which may impact their early speech production skills. Indeed, Lachs, Pisoni, and Kirk (2001) recently found that deaf children who were better at integrating auditory and visual information of spoken words in sentences produced more intelligible speech. However, the participants in the Lachs et al. study were much older children than the deaf infants we are studying now. Investigating the audiovisual speech perception skills of deaf infants who receive CIs would provide valuable information about how providing CIs at an early age may affect audiovisual integration skills and may be useful for predicting their later speech and language production skills (Bergeson & Pisoni, in press).

Conclusions

We have adapted the VH procedure to assess the speech perception skills of deaf infants who have received CIs. So far, the results are very encouraging. The attrition rates are relatively low, and deaf infants who have received CIs are showing trends in their looking times that are similar to findings obtained with NH 6-month-olds. This pattern of responses was observed strongly in one participant who received his CI at 6 months of age and was studied repeatedly over time. We suspect that earlier implantation may facilitate the development of attention to speech sounds because sound and the information specified by sound sources in the environment will become available at an earlier point in neural and perceptual development. Attention to speech and spoken language is an important prerequisite for learning about the organization of sounds in the ambient language and developing knowledge of the sound patterns and regularities of sounds. The initial results obtained so far using the VH procedure are consistent with the general hypothesis that early exposure to sound and especially exposure to speech underlies the development of auditory attention and speech discrimination skills, although more data from more deaf and NH infants will be needed to see if these trends are reliable.

References

Bergeson, T.R., & Pisoni, D.B. (in press). Audiovisual speech perception in deaf adults and children following cochlear implantation. In G. Calvert, C. Spence, & B.E. Stein (Eds.), *Handbook of Multisensory Integration.*

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P.W., Kennedy, L.J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General, 117*, 21-33.

Best, C.T. (1995). Learning to perceive the sound patterns of English. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in Infancy Research* (Vol. 9, pp. 217-304). Norwood, NJ: Ablex.

Best, C.T., McRoberts, G.W., & Sithole, N.M. (1988). Examination of the perceptual re-organization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

Bolinger, D.L., & Gerstman, L.J. (1957). Disjuncture as a cue to constraints. *Word, 13*, 246-255.

Boysson-Bardies, B.de, Vihman, M.M., Roug-Hellichius, L., Durand, C., Landberg, I., & Arao, F. (1992). Material evidence of selection from the target language. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications.* Timonium, Md.: York Press.

Church, K. (1987). Phonological parsing and lexical retrieval. *Cognition, 25*, 53-69.

Cohen, L.B., Atkinson, D.J., & Chaput, H.H. (2000). *Habit 2000: A new program for testing infant perception and cognition*. (Version 1.0). Austin: University of Texas.

Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133-142.

Echols, C.H., Crowhurst, M.J., & Childers, J. (1997). Perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language, 36*, 202-225.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*(968), 303-306.

Friederici, A.D., & Wessels, J.M.I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception & Psychophysics, 54*, 287-295.

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

Gaver, W.W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology, 5,* 1-29.

Goodsitt, J.V., Morgan, J.L., & Kuhl, P.K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language, 20*, 229-252.

Hayes, D. & Northern, J.L. (1996). *Infants and hearing*. San Diego, CA: Singular.

Hillenbrand, J.M., Minifie, F.D., & Edwards, T.J. (1979). Tempo of spectrum change as a cue in speech sound discrimination by infants. *Journal of Speech and Hearing Research, 22*, 147-165.

Horowitz, F.D. (1975). Infant attention and discrimination: Methodological and substantive issues. *Monographs of the Society for Research in Child Development, 39*(Serial no. 158).

Houston, D.M., Carter, A.K., Pisoni, D.B., Kirk, K.I., & Ying, E.A. (in press). Word learning skills of deaf children following cochlear implantation: A first report. In *Research on Spoken Language Processing Progress Report No. 25*. Bloomington, IN: Speech Research Laboratory, Indiana University.

Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*, 548-567.

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P.W. (2002). How infants adapt speech-processing capacities to native-language structure. *Current Directions in Psychological Science, 11*, 18.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1-23.

Jusczyk, P.W., Copan, H., & Thompson, E. (1978). Perception by 2-month-old infants of glide contrasts in multisyllabic utterances. *Perception & Psychophysics, 24*, 515-520.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675-687.
Jusczyk, P.W., Hohne, E.A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics, 61*, 1465-1476.
Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159-207.
Jusczyk, P.W., Luce, P.A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*, 630-645.
Kirk, K.I., Miyamoto, R.T., Ying, E.A., Perdew, A.E., & Zuganelis. (in press). Cochlear implantation in young children: Effects of age at implantation and communication mode. *Volta Review*.
Kral, A., Hartmann, R., Tillein, J., Held, S., & Klinke, R. (2000). Congenital auditory deprivation reduces synaptic activity within the auditory cortex in a layer-specific manner. *Cerebral Cortex, 10*, 714-726.
Kuhl, P.K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development, 6*, 263-285.
Kuhl, P.K., & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science, 218,* 1138-1141.
Kujala, T., Alho, K., & Naatanen, R. (2000). Cross-modal reorganization of human cortical functions. *Trends in Neuroscience, 23*, 115-120.
Lachs, L., Pisoni, D.B., & Kirk, K.I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear & Hearing, 22,* 236-251.
Leake, P.A., & Hradek, G.T. (1988). Cochlear pathology of long term neomycin induced deafness in cats. *Hearing Research, 33*, 11-34.
Levitt, A.G., Utman, J., & Aydelott, J. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language, 19,* 19-49.
Mattys, S.L., & Jusczyk, P.W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition, 78*, 91-121.
Mattys, S.L., Jusczyk, P.W., Luce, P.A., & Morgan, J.L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38*, 465-494.
Miller, J.L., & Eimas, P.D. (1983). Studies on the categorization of speech by infants. *Cognition, 13*, 135-165.
Moffitt, A.R. (1971). Consonant cue perception by twenty-to-twenty-four-week old infants. *Child Development, 42*, 717-731.
Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development, 66*, 911-936.
Neville, H.J., & Bruer, J.T. (2001). Language processing: How experience affects brain organization. In J.D.B. Bailey & J.T. Bruer (Eds.), *Critical thinking about critical periods* (pp. 151-172). Baltimore: Paul H. Brookes.
Polka, L., & Werker, J.F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.
Ponton, C.W., & Eggermont, J.J. (2001). Of kittens and kids: Altered cortical maturation following profound deafness and cochlear implant use. *Audiology & Neuro-Otology, 6*, 363-380.
Posner, M.I., & Rothbart, M.K. (2000). Developing mechanisms of self-regulation. *Development & Psychopathology, 12*, 427-441.
Rauschecker, J.P., & Korte, M. (1993). Auditory compensation for early blindness in cat cerebral cortex. *Journal of Neuroscience, 13*, 4538-4548.
Rebscher, S.J., Snyder, R.L., & Leake, P.A. (2001). The effect of electrode configuration and duration of deafness on threshold and selectivity of responses to intracochlear electrical stimulation. *Journal of the Acoustical Society of America, 109*, 2035-2048.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Shepherd, R.K., & Hardie, N.A. (2001). Deafness-induced changes in the auditory pathway: Implications for cochlear implants. *Audiology & Neuro-Otology, 6*, 305-318.

Suomi, K. (1993). An outline of a developmental model of adult phonological organization and behavior. *Journal of Phonetics, 21*, 29-60.

Svirsky, M.A., Stallings, L.M., Lento, C.L., Ying, E., & Leonard, L.B. (2002). Grammatical morphologic development in pediatric cochlear implant users may be affected by the perceptual prominence of the relevant markers. *Annals of Otology, Rhinology & Laryngology Supplement 2002; 111(Suppl 198),* 109-112

Tobey, E., Pancamo, S., Staller, S., Brimacombe, J., & Beiter, A. (1991). Consonant production in children receiving a multichannel cochlear implant. *Ear & Hearing, 12*, 23-31.

Trehub, S.E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology, 9*, 91-96.

Trehub, S.E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47*, 466-472.

Tsushima, T., Takizawa, O., Sasaki, M., Siraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994, October). *Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: Language specific developmental changes in speech perception abilities.* Paper presented at the International Conference on Spoken Language Processing, Yokohama, Japan.

van de Weijer, J. (1998). *Language input for word discovery* (Vol. 9). Nijmegen, The Netherlands: Max Planck Institute.

Waltzman, S.B., & Cohen, N.L. (1998). Cochlear implantation in children younger than 2 years old. *The American Journal of Otology, 19*, 158-162.

Werker, J.F., & Lalonde, C.E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*, 672-683.

Werker, J.F., Shi, R., Desjardins, R., Pegg, J.E., Polka, L., & Patterson, M. (1998). Three methods for testing infant speech perception. In A. Slater (Ed.), *Perceptual development: Visual, auditory, and speech perception in infancy.* East Sussex, UK: Psychology Press.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49-63.

Woodward, J.Z., & Aslin, R.N. (1990, April). *Segmentation cues in maternal speech to infants.* Paper presented at the 7th biennial meeting of the International Conference on Infant Studies, Montreal, Quebec, Canada.

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

Cochlear Implants: Signal Processing and Speech Perception[1]

Philipos C. Loizou[2] and Mario A. Svirsky[3]

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

[1] This work was supported in part by Grants No. R01 DC03421, R01 DC03739, R01 DC 00421 and R01 DC00064 from the National Institute of Deafness and other Communication Disorders, NIH.
[2] The University of Texas-Dallas.
[3] DeVault Otologic Research Laboratory, Indiana University School of Medicine, Indianapolis, IN.

Cochlear Implants: Signal Processing and Speech Perception

Abstract. Several million Americans today have profound hearing loss, and for years they had to rely on conventional hearing aids. Although hearing aids have been found to benefit hearing impaired individuals who suffer mild to moderate deafness, they do not seem to provide much benefit to individuals with profound (sensorineural) deafness. Today, a prosthetic device, a *cochlear implant*, can be implanted in the inner ear to restore partial hearing to profoundly deaf people. Cochlear implants are now established as a new option for individuals with profound hearing impairments. Most individuals who are implanted with cochlear prosthesis are able to understand speech without lip-reading and can communicate over the phone. This article presents an overview of cochlear implants, describing the signal processing and stimulation strategies they employ and summarizing key research findings.

Background

Hearing and Deafness.

In a healthy ear, sound undergoes a series of transformations as it travels through the outer ear, middle ear, inner ear, auditory nerve and into the brain. The outer ear picks up acoustic pressure waves, which are converted to mechanical vibrations by a series of small bones in the middle ear. In the inner ear, the cochlea, a snail-shaped cavity filled with fluid, transforms the mechanical vibrations to vibrations in fluid. Pressure variations within the fluid of the cochlea lead to displacements of a flexible membrane, called the basilar membrane. These displacements contain information about the frequency of the acoustic signal. Attached to the basilar membrane are hair cells, which are bent according to the displacements of the basilar membrane. The bending of the hairs releases a substance that causes neurons to fire, signaling the presence of excitation at a particular site in the inner ear. These neurons communicate with the central nervous system and transmit information about the acoustic signal to the brain.

The hair cells in conjunction with the basilar membrane are responsible for translating mechanical information into neural information. If the hair cells are damaged, the auditory system has no way of transforming acoustic pressure waves (sound) to neural impulses, resulting in hearing impairment. The hair cells can be damaged by certain diseases (e.g., meningitis, Meniere's disease), congenital disorders, drug treatments, or by many other causes. Damaged hair cells can subsequently lead to degeneration of adjacent auditory neurons. If a large number of hair cells or auditory neurons are damaged, then the condition is called profound deafness. Research (e.g., Hinojosa & Marion, 1983) has shown that the most common cause of deafness is the loss of hair cells rather than the loss of auditory neurons. This was very encouraging for cochlear implantation because remaining neurons could be excited directly through electrical stimulation. A cochlear prosthesis therefore bypasses the normal hearing mechanism (outer, middle, and part of the inner ear, including the hair cells) and electrically stimulates the remaining auditory neurons directly. The challenge we face is finding the optimal signal processing to stimulate (electrically) auditory neurons so that meaningful information about speech is conveyed to the brain. For example, information about the amplitude and the frequency of the acoustic signal should be conveyed.

Encoding Frequency.

The question then arises: "How does the auditory system encode frequencies?" The pioneering work of Georg von Bekesy in the 1950’s showed that the basilar membrane in the inner ear is responsible for analyzing the input signal into different frequencies. Different frequencies cause maximum vibration amplitude at different points along the basilar membrane (see Fig. 1). Low frequency sounds create traveling waves in the fluids of the cochlea, which cause the basilar membrane to vibrate, with largest amplitude of displacement at the apex (see Fig. 1) of the basilar membrane. On the other hand, high frequency sounds create traveling waves with largest displacement at the base of the basilar membrane (near the stapes). If the signal is composed of multiple frequencies, then the resulting traveling wave will create maximum displacement at different points along the basilar membrane. The cochlea thereby acts as a spectrum analyzer, decomposing complex sounds into their frequency components.

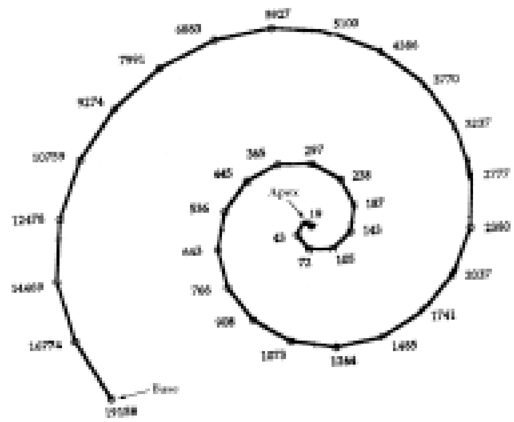


Figure 1. Diagram of the basilar membrane showing the base and the apex. The position of maximum displacement in response to sinusoids of different frequency (in Hz) is indicated.

The corresponding hair cells bent by the displacement in the membrane stimulate adjacent nerve fibers, which are organized according to the frequency at which they are most sensitive. Each place or location in the cochlea is therefore responding "best" to a particular frequency. This mechanism for determining frequency is referred to as *place coding*. The place mechanism for coding frequencies has motivated multi-channel cochlear implants.

Cochlear Implants.

Cochlear implants are based on the premise that there are sufficient auditory nerve fibers remaining for stimulation in the vicinity of the electrodes. Once the nerve fibers are stimulated, they fire and propagate neural impulses to the brain. The brain interprets these impulses as sounds. The perceived loudness of the sound may depend on the number of nerve fibers activated and their rates of firing. If a large number of nerve fibers is activated, then the sound is perceived as loud and vice versa. The number of fibers activated is a function of the amplitude of the stimulus current. The loudness of the sound can therefore be controlled by varying the amplitude of the stimulus current. The pitch, on the other hand, is related to the place in the cochlea that is being stimulated. Low pitch sensations are elicited when electrodes near the apex are stimulated, while high pitch sensations are elicited by stimulation of electrodes near the base. Thus, the implant can effectively transmit information to the brain about the loudness of the sound that is a function of the amplitude of the stimulus current and the sound pitch that is a function of the place in the cochlea being stimulated. Additionally, the rate at which electrical pulses are delivered also affects the perceived pitch.

Several cochlear implant devices have been developed over the years (Wilson, 1993; Loizou, 1998). All of the implant devices have the following features in common: a microphone that picks up the sound, a signal processor that converts the sound into electrical signals, a transmission system that transmits the electrical signals to the implanted electrodes, and an electrode or an electrode array (consisting of multiple electrodes) that is implanted into the cochlea. In single-channel implants (which are now considered obsolete, due to the inferior perceptual results they provide) only one electrode is used. In multi-channel cochlear implants, an electrode array is inserted in the cochlea so that different auditory nerve fibers can be stimulated at different places, thereby exploiting the place mechanism for coding frequencies. Different electrodes are stimulated, depending on the frequency of the signal. Electrodes near the base of the cochlea are stimulated with high frequency signals, while electrodes near the apex are stimulated with low frequency signals. The signal processor is responsible for breaking the input signal into different frequency bands or channels, and delivering the filtered signals to the appropriate electrodes. The main function of the signal processor is to decompose the input signal into its frequency components, much like a healthy cochlea. The designers of cochlear prosthesis are faced with the challenge of developing signal processing techniques that mimic the function of a healthy cochlea.

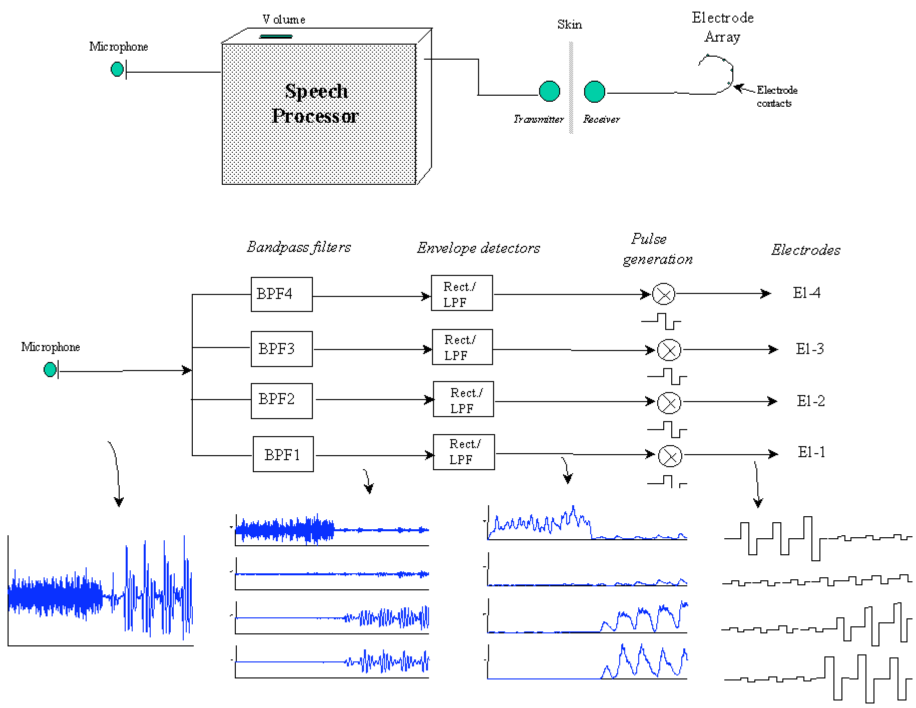


Figure 2. Diagram showing the operation of a four-channel cochlear implant. Sound is picked up by a microphone and sent to a speech processor box worn by the patient. The sound is then processed, and electrical stimuli are delivered to the electrodes through a radio-frequency link. Bottom figure shows a simplified implementation of the CIS signal processing strategy using the syllable "sa." The signal first goes through a set of four bandpass filters that divide the acoustic waveform into four channels. The envelopes of the bandpassed waveforms are then detected by rectification and low-pass filtering. Current pulses are generated with amplitudes proportional to the envelopes of each channel and transmitted to the four electrodes through a radio-frequency link. Note that in the actual implementation, the envelopes are compressed to fit the patient's electrical dynamic range.

Figure 2 shows, as an example, the operation of a four-channel implant. Although the example uses four channels for explanatory purposes, cochlear implants currently in use typically employ between eight and twenty channels. Sound is picked up by a microphone and sent to a speech processor box worn by the patient. Speech processors can be body-worn or behind-the-ear (BTE). In the case of some devices, the BTE version is less flexible than the body-worn processor. After the sound is picked up by the microphone, it is processed through a set of four bandpass filters, which divide the acoustic signal into four channels. Current pulses are generated with amplitudes proportional to the energy in each channel, and transmitted to the four electrodes through a radio-frequency link. The relative amplitudes of the current pulses delivered to the electrodes reflect the spectral content of the input signal (Fig. 2). For instance, if the speech signal contains mostly high frequency information (e.g., /s/), then the pulse amplitude of channel 4 will be large relative to the pulse amplitudes of channels 1-3. Similarly, if the speech signal contains mostly low frequency information (e.g., vowel /a/) then the pulse amplitude of channels 1 and 2 will be large relative to the amplitudes of channels 3 and 4 (Fig. 2).

Who can be implanted?

Not all people with hearing impairment are candidates for cochlear implantation. Certain audiological criteria need to be met. First, the hearing loss has to be severe or profound and it must be bilateral. Hearing loss is typically measured as the average of pure tone hearing thresholds at 500, 1000 and 2000 Hz, expressed in dB with reference to normal thresholds. Profound deafness is defined as a hearing loss of 90 dB or more, and severe deafness is a hearing loss of 70 dB or more. Second, the candidate should not derive substantial benefit from hearing aids. For adults, this means that they must have acoustically aided scores of 50% or less with the ear to be implanted, and 60% or less with the unimplanted ear (or bilaterally), in a sentence recognition test such as the Hearing-In-Noise-Test (HINT) (Nilsson, Soli & Sullivan, 1994). Typically, the test is administered with no visual cues and in quiet conditions. Children age two years or older with bilateral, severe to profound sensorineural loss are also candidates for cochlear implantation if they receive little or no useful benefit from hearing aids and show a lack of progress in the development of auditory skills. Children between 12 and 24 months of age may also be implanted if they have profound bilateral deafness and do not show progress in the development of auditory skills. In special cases, implantation may be indicated even before 12 months of age. One case is deafness due to meningitis, which may cause ossification of the cochlea, making a delayed implantation more difficult and potentially less successful.

Implant Characteristics.

Commercially available implant devices differ in the following characteristics: Electrode design (e.g., number of electrodes, electrode configuration); type of stimulation (e.g., analog or pulsatile); and signal processing (e.g., waveform representation, envelope representation or spectral features). A brief description of each of the above characteristics is given below.

Electrode Design.

Electrodes are commonly placed in the scala tympani because it brings them in close proximity with auditory neurons that lie along the length of the cochlea. This electrode placement is preferred because it preserves the "place" mechanism of the normal cochlea for coding frequencies. That is, auditory neurons that are "tuned" for high frequencies are stimulated whenever the electrodes near the base are stimulated, whereas auditory neurons that are "tuned" for low frequencies are stimulated whenever the electrodes near the apex are stimulated. In most cases, the electrode arrays can be inserted in the scala tympani to depths of 22-30 mm within the cochlea.

The number of electrodes, as well as the spacing between them, affects the place resolution for coding frequencies. In principle, the larger the number of electrodes, the finer the place resolution for coding. Frequency coding is constrained, however, by two inherent factors: (1) number of surviving auditory neurons that can be stimulated at a particular site in the cochlea, and (2) spread of excitation associated with electrical stimulation. Unfortunately, not much can be done about the first problem, because it depends on the etiology of deafness. Ideally, we would like to have surviving auditory neurons lying along the length of the cochlea. Such a neuron survival pattern would support good frequency representation through the use of multiple electrodes, each stimulating a different site in the cochlea. At the other extreme, consider the situation where the number of surviving auditory neurons is restricted to a small area in the cochlea. In that situation, a few electrodes implanted near that area would be as good as 100 electrodes distributed along the cochlea. So, using a large number of electrodes will not necessarily result in better performance, because frequency coding is constrained by the number of surviving auditory neurons that can be stimulated.

In addition, frequency coding is constrained by the spread of excitation caused by electrical stimulation. Electric current injected into the cochlea tends to spread out symmetrically from the source. As a result, the current does not stimulate just a single (isolated) site of auditory neurons but several. Such a spread in excitation is most prominent in the monopolar electrode configuration. In this configuration, the active electrode is located far from the reference electrode, which acts as a ground for all electrodes. The spread of excitation can be constrained, to a degree, by using a bipolar electrode configuration. In this configuration, the active and the reference (ground) electrodes are placed close to each other. Bipolar electrodes have been shown to produce a more localized stimulation than monopolar (van den Honert & Stypulkowski, 1987; Merzenich & White, 1977). Although the patterns of electrical stimulation produced by these two configurations are different, it is still not clear which will result in better performance for a particular patient.

Currently, some implant devices employ monopolar electrodes, other devices employ bipolar electrodes, and yet other devices support both types. Table 1 lists some current implant devices and their characteristics. The Nucleus device uses 22 electrodes spaced 0.75 mm apart. Electrodes that are 1.5 mm apart are used as bipolar pairs. The Clarion device provides both monopolar and bipolar configurations, with 8 electrodes spaced 2 mm apart. The Med-El device uses 12 electrodes in monopolar configuration.

Device	Number of electrodes	Electrode configuration	Type of stimulation
Nucleus	22	Bipolar	pulsatile
Clarion	8	Monopolar/Bipolar	analog/pulsatile
Med-El	12	Monopolar	pulsatile

Table 1: Characteristics of commercially available cochlear implant devices.

Some companies have developed “modiolar-hugging” electrodes as an upgrade or an alternative to their standard electrodes. Modiolar-hugging electrodes are located closer to the inner wall of the cochlea and therefore are closer to the neural elements that receive electrical stimulation (Cohen et al., in press). It has been demonstrated that this electrode design results in lower perceptual thresholds and comfortable levels, which in turn decreases power consumption and therefore increases battery life (Saunders et al., in press; Parkinson et al., in press). Another potential benefit of modiolar-hugging electrodes is that their stimulation may be more focused on smaller groups of neurons. This may result in better discrimination of place of stimulation in the cochlea (or, equivalently, frequency discrimination) and in improved speech perception. However, this potential benefit has not yet been demonstrated in a clear way.

Type of Stimulation.

Information is presented either in analog or pulsatile form. In analog stimulation, an electrical analog of the acoustic waveform is presented to the electrode. In multi-channel implants, the acoustic waveform is bandpass filtered, and the filtered waveforms are presented to all electrodes simultaneously. The rationale behind this type of stimulation is that the nervous system will sort out and/or make use of all the information contained in the raw acoustic waveforms. One disadvantage of analog stimulation is that its simultaneous action may cause channel interactions. The SAS strategy implemented in the Clarion device is the only strategy that uses analog stimulation.

In pulsatile stimulation, the information is delivered to the electrodes using a set of narrow pulses. In some devices, the amplitudes of these pulses are extracted from the envelopes of the filtered waveforms (Fig. 2). The advantage of this type of stimulation is that the pulses can be delivered in a non-overlapping (i.e., non-simultaneous) fashion, thereby minimizing channel interactions. The majority of the commercial implant devices utilize pulsatile stimulation.

Information is transmitted to the electrodes using a transcutaneous connection, where an external transmitter encodes the stimulus information for radio-frequency transmission from an external coil to an implanted coil. The internal receiver decodes the signal and delivers the stimuli to the electrodes. All cochlear implant devices (e.g., Nucleus, Clarion, Med-El) today use transcutaneous connections. Implanted receivers use a small magnet to facilitate alignment of the external coil to the implanted coil. Although the use of this magnet is well justified, it does present compatibility problems with MRI scanners.

Signal Processing.

The last, and perhaps most important, difference among implant devices is in the signal processing strategy used for transforming the speech signal to electrical stimuli. Some of these strategies are aimed at preserving waveform information, others are aimed at preserving envelope information, and yet others aimed at preserving spectral features (e.g., formants). A more detailed discussion on these signal-processing techniques can be found in Loizou (1998).

Speech Processing Strategies.

Several speech-processing strategies have been developed over the years for multi-channel implants (Loizou, 1998). The effectiveness and performance of these speech processing strategies improved significantly over the years. The Nucleus multi-channel implant, for instance, underwent several changes before developing the latest speech processing strategy, the Advanced Combined Encoder (ACE) strategy. Below, we give the description of the three most common and most successful speech processing strategies used commercially today: the Continuous Interleaved Sampling (CIS), the Simultaneous Analog Stimulation (SAS) and the ACE/SPEAK/n-of-m strategies. The CIS strategy is available in all three implant devices, the SAS is available only in the Clarion device and the ACE/SPEAK/n-of-m strategy is available in the Nucleus-24 and Med-El devices.

Continuous Interleaved Sampling (CIS) Strategy.

The CIS strategy was developed by researchers at the Research Triangle Institute (Wilson et al., 1991). The block diagram of the CIS strategy is shown in Figure 3 (which is a more complete version of the block diagram shown in Figure 2). The signal is first pre-emphasized and then passed through a bank of n bandpass filters, where n typically corresponds to the number of electrodes. The envelopes of the n filtered waveforms are extracted by full-wave rectification and low-pass filtering. The envelope outputs are finally compressed and then used to modulate biphasic pulses. A logarithmic compression function is used to ensure that the envelope outputs fit the patient's dynamic range of electrically evoked hearing. Trains of balanced biphasic pulses, with amplitudes proportional to the envelopes, are delivered to the six electrodes at a constant rate in a non-overlapping fashion i.e., such that only one electrode is stimulated at a time. The rate at which the pulses are delivered to the electrodes has been found to influence speech recognition. High pulse-rate stimulation (i.e., 800 pulses per second, per channel, or higher) typically yields higher performance than low stimulation rate (e.g., Loizou, Poroy & Dorman, 2000). There are also other CIS parameters that may affect performance, and these are discussed in a later section. The CIS strategy is currently being used in the Med-El, the Clarion and the Nucleus CI24M devices. It should be pointed out that the CIS strategy is implemented differently in the various devices, and this can account for some of the differences in performance obtained with the CIS strategy in the commercial devices.

Figure 4 shows, as an example, the pulsatile waveforms produced for the syllable "sa" using a simplified, four-channel implementation of the CIS strategy.

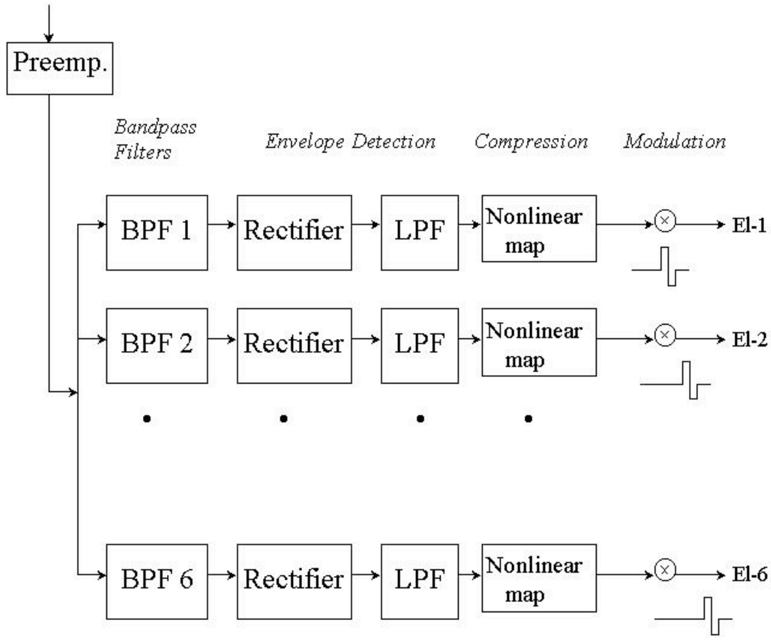


Figure 3. Block diagram of the CIS strategy. The signal is first pre-emphasized and filtered into six frequency bands. The envelopes of the filtered waveforms are then extracted by full-wave rectification and low-pass filtering. The envelope outputs are compressed to fit the patient's dynamic range and then modulated with biphasic pulses. The biphasic pulses are transmitted to the electrodes in an interleaved fashion

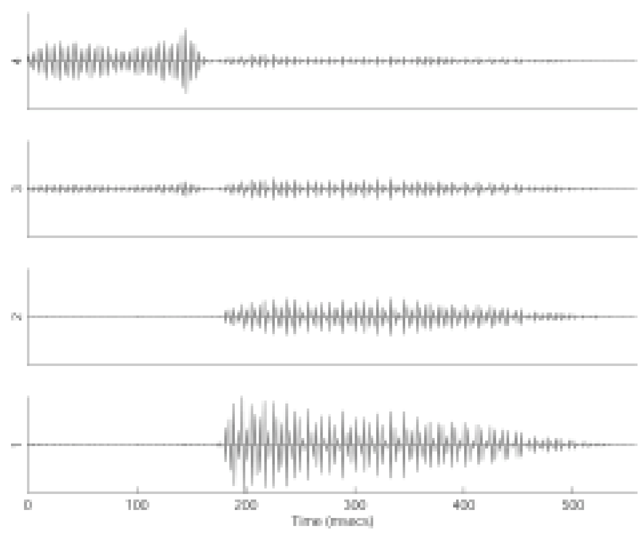


Figure 4. Pulsatile waveforms of the syllable "sa" produced by a simplified implementation of the CIS strategy using a 4-channel implant. The pulse amplitudes reflect the envelopes of the bandpass outputs for each channel. The pulsatile waveforms are shown prior to compression.

In the CIS strategy, the electrodes are stimulated in a non-overlapping manner, i.e., not simultaneously. This is done to address a major concern associated with simultaneous stimulation, which is channel interaction. This interaction is caused by the summation of electrical fields from individual electrodes, particularly when the electrodes are configured in monopolar mode. Neural responses to stimuli from one electrode may be significantly distorted by stimuli from other electrodes. These interactions may distort speech spectrum information and therefore degrade speech understanding.

ACE/SPEAK/N-of-M Strategy.

Unlike previous strategies developed for the Nucleus implant, the SPEAK strategy does not extract any features (e.g., F1, F2) from the speech waveform. Instead it analyzes the speech signal using a bank of 20 bandpass filters and a spectral maxima detector. The signal from the microphone is first pre-amplified and then sent through a bank of 20 bandpass filters with center frequencies ranging from 250 Hz to 10 kHz. The output of each filter is rectified and low-pass filtered with a cutoff frequency of 200 Hz. The SPEAK processor continuously estimates the outputs of the 20 filters and selects the ones with the largest amplitude. The number of maxima selected varies from 5 to 10, depending on the spectral composition of the input signal, with an average number of 6 maxima. The selected electrodes are stimulated at a rate that varies between 180 and 300 Hz depending on: (1) the number of maxima selected and (2) on the patient's individual parameters. The selected amplitudes of the spectral maxima are then logarithmically compressed, to fit the patient's electrical dynamic range, and transmitted to the selected electrodes through a radio-frequency link.

One electrode is allocated for each of the 20 filter outputs, according to the tonotopic order of the cochlea. That is, the most apical electrode is allocated to the filter with the lowest center frequency, while the most basal electrode is allocated to the filter with the highest center frequency. Figure 5 illustrates the pattern of electrical stimulation for the word "choice." As can be seen, the electrodes selected for stimulation in each cycle vary depending upon the spectral content of the signal.

The SPEAK strategy is available in the Nucleus CI24M device. A strategy called n-of-m that is similar to the SPEAK strategy is available in the Med-El device. In the n-of-m strategy, the number of maxima selected in each cycle is fixed. Therefore, “n” corresponds to the number of maxima selected, and “m” corresponds to the total number of channels available. The SPEAK strategy is therefore similar (but not exactly identical) to a 6-of-20 strategy, with an average stimulation rate of 250 cycles per second.

The ACE strategy, described in detail by Vandali et al. (2000) is the latest strategy available for the Nucleus device. It is essentially similar to the n-of-m strategy although some implementation details do differ. For example, the frequency analysis is done using a 128 Hanning window and FFT rather than using an actual software filter bank. Another implementation detail of the ACE strategy is that the FFT’s are calculated at a rate of 760 times per second or less. When the stimulation rate is less than 760 cycles per second, the FFT analysis rate is set to equal the stimulation rate, as is normally done in all other DSP-based strategies used by cochlear implant speech processors. However, when the stimulation rate exceeds 760 cycles per second, the analysis rate is limited to 760 times per second and the higher stimulation rate is achieved by repeating stimulation frames as necessary. This compromise is not likely to affect the speech perception of ACE users because the envelopes of the signals coming out of the filterbank change at rates that are typically much lower than 760 Hz (and indeed, as discussed above, these envelopes are usually lowpass filtered using cutoff frequencies of a few hundred Hz at most).

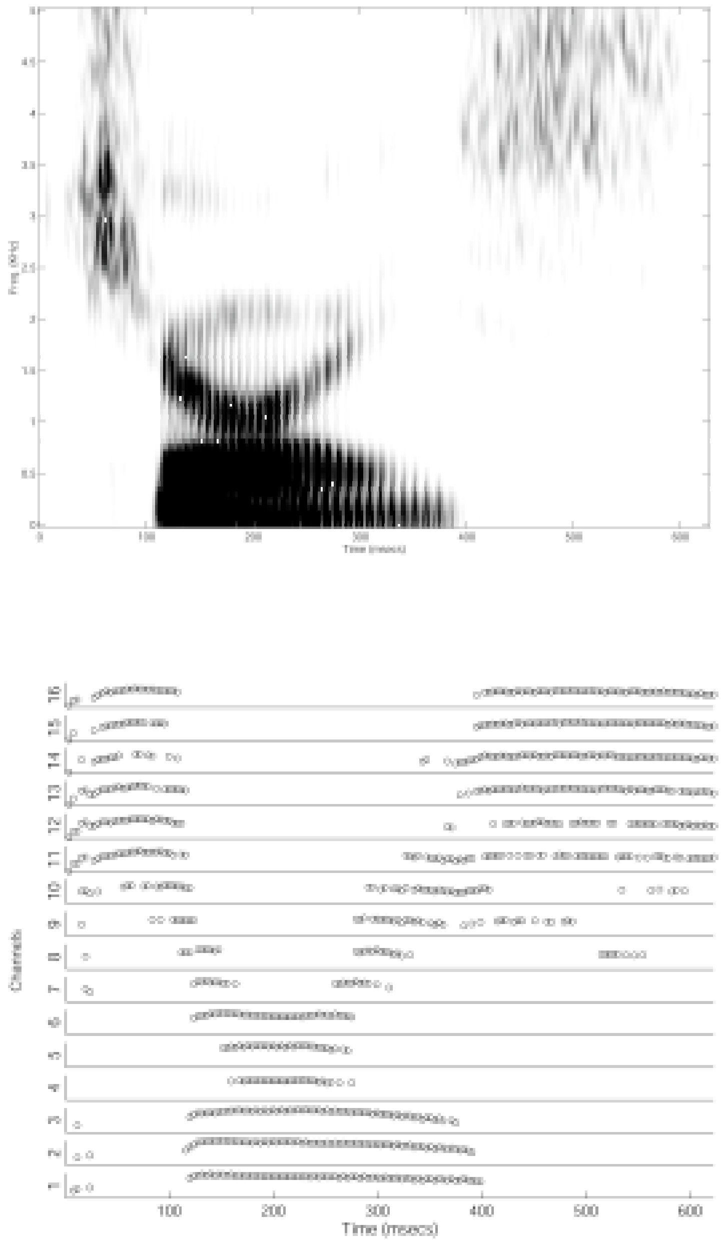


Figure 5. Example of the SPEAK strategy using the word "choice." The top panel shows the spectrogram of the word "choice," and the bottom panel shows the filter outputs selected at each cycle. The channels selected for stimulation depend upon the spectral content of the signal. As shown in the bottom panel, during the "s" portion of the word, high frequency channels (10-16) are selected and during the "o" portion of the word, low frequency channels (1-6) are selected.

Simultaneous Analog Stimulation (SAS) Strategy.

In the SAS strategy (Kessler, 1999), which is available only in the Clarion device, the signal is first pre-emphasized and then passed through a bank of seven bandpass filters. The seven filtered waveforms are compressed to fit the patient's dynamic range and then used to stimulate seven electrodes simultaneously. Pseudo-analog waveforms are delivered to each electrode at a rate of 13,000 samples/sec per channel. To minimize possible channel interaction, the electrodes are configured in bipolar mode that provides a more selective electrical stimulation pattern than monopolar coupling.

Importance of Fitting: Optimizing Patient Performance.

Cochlear implant manufacturers now offer a multitude of speech processing strategies in their speech processors. It is generally not known, however, which strategy will work the best for a particular patient. As a result, clinicians now have the option to program the patient with multiple strategies and have the patient select the strategy they prefer. In addition, clinicians have the capability, thanks to the flexible fitting software, to change certain speech processing parameters to optimize performance.

For patients fitted with the CIS strategy, clinicians can vary a number of parameters to optimize speech recognition performance for each patient. These parameters include pulse rate, pulse duration and stimulation order.

Pulse Rate and Pulse Duration.

The pulse rate defines the number of pulses per sec (pps) delivered to each electrode. Pulse rates as low as 100 pulses/sec and as high as 2400 pulses/sec have been used commercially. The "optimal" pulse rate for speech recognition varies from patient to patient. Wilson et al. (1995) reported that some patients obtain a maximum performance on the 16-consonant recognition task with 833 pulses/sec and pulse duration of 33 msecs/phase. Other patients obtained maximum performance at different combinations of pulse rate and pulse duration (Wilson, Lawson & Zerbi, 1993). Loizou et al. (2000) showed that the performance of some patients on word recognition increased monotonically from 400 pps to 2100 pps. The performance obtained at 2100 pps and with a 40-ms pulse duration, was found to be significantly higher than the performance obtained at 800 pps. However, note that studies conducted with users of the ACE strategy did not find a similar, consistent advantage for higher stimulation rates. Vandali et al. (2000) found that the use of rates higher than 250 cycles per second (up to 1615 cps) did not provide significant improvement in speech comprehension. Holden, Skinner and Holden (in press) found that higher rates resulted in significantly better performance for some subjects and significantly worse performance for other individuals.

Stimulation Order.

The stimulation order refers to the order that electrodes are stimulated. The stimulation order can be varied to minimize possible interaction between channels. One possibility is to stimulate the electrodes in an apex-to-base order, i.e., first stimulate electrode 1, then 2, etc., and lastly, 6. This way, signals in the low frequencies (apex) are stimulated first, and signals in the high frequencies (base) are stimulated last. This apex-to-base order, however, does not minimize the spatial separation between sequentially stimulated electrodes. Alternatively, the electrodes can be stimulated in a so-called "staggered" order, i.e., 6-3-5-2-4-1, which maximizes the spatial separation between stimulated electrodes. As with the pulse rate, preference for stimulation order varies from patient to patient. Some patients prefer the apex-to-base stimulation because they say speech sounds more natural and intelligible while other patients prefer the staggered order stimulation.

For patients fitted with the Nucleus' CI24M device, clinicians have the option to select a subset of electrodes for stimulation. Research has shown that stimulating only a subset of electrodes, rather than all 22 electrodes, can produce significant benefits for some patients fitted with the CIS strategy. Plant et al. (1999) for instance, showed that some subjects preferred and performed better on word recognition with the CIS strategy when it was programmed with 8 channels than when it was programmed with 16 channels.

For patients fitted with the Clarion device, clinicians have the option to fit the patient with the CIS, the SAS or the Paired Pulsatile Stimulation (PPS) strategies. The PPS strategy is similar to CIS with the exception that two simultaneous pulses are delivered at a time instead of just one. Research has shown that some patients prefer the SAS strategy to the CIS strategy (Osberger & Fisher, 1999; Battmer, Zilberman, Haake & Lenarz, 1999), while other patients prefer the PPS strategy to the CIS strategy (Armstrong-Bendall et al., 1999). In most cases, Clarion patients seem to perform better on speech recognition tasks with the strategy they prefer.

For patients fitted with the Med-El device (COMBI 40+), clinicians now have the option to fit patients with the n-of-m strategy in addition to the CIS strategy. Research has shown that Med-El patients performed better with the 7-of-12 strategy (operating at a higher rate) than the CIS strategy (12 channels) on monosyllabic word recognition (Ziese et al., 2000). This is consistent with research reported by Brill et al. (1997) showing that by trading channels with higher simulation rates (i.e., by using fewer number of channels to obtain higher rates of stimulation) performance can be improved.

Factors Affecting the Performance of Cochlear Implant Patients.

There is a great variability in the speech recognition performance of cochlear implant patients. For a given type of implant, auditory performance may vary from zero to nearly 100% correct. Auditory performance is defined here as the ability to discriminate, detect, identify, or recognize speech (a typical measure of auditory performance is the percent correct score on open-set speech recognition tests). The factors responsible for such variability in auditory performance have been the focus of research for many years (Blamey et al., 1996). Some of the factors that have been found to affect auditory performance are listed below.

Duration of Deafness.

The duration of deafness prior to implantation has been found to have a negative effect on auditory performance. Individuals with shorter durations of auditory deprivation tend to achieve better auditory performance than individuals with longer durations of auditory deprivation.

Age of Onset of Deafness.

The age of onset of deafness has a major impact on the success of cochlear implants, depending on whether the deafness was acquired before (prelingual) or after (postlingual) learning speech and language. It is now well established that children and adults with postlingual deafness perform better than those with prelingual or congenital deafness. However, this factor interacts with age at implantation, because congenitally deaf children can achieve excellent communication skills if they are implanted early enough. This includes speech perception (Meyer, Svirsky, Kirk & Miyamoto, 1998; Svirsky & Meyer, 1999; Meyer & Svirsky, 2000), speech production (Svirsky, Sloan, Caldwell & Miyamoto, 2000b) and language development (Svirsky, Robbins, Kirk, Pisoni & Miyamoto, 2000a).

Age at Implantation.

Prelingually deafened persons who were implanted in adolescence have been found to obtain higher levels of auditory performance than those implanted in adulthood. People implanted at an early age seem to perform better than people implanted in adulthood. In addition, recent results suggest that congenitally deaf children who are implanted earlier in life have a better communicative prognosis (Svirsky et al., 2000a).

Duration of Cochlear Implant Use.

Duration of experience with the implant has been found to have a strong positive effect on auditory performance for both adults and children.

Other Factors.

Other factors that may affect auditory performance include: (1) number of surviving spiral ganglion cells, (2) electrode placement and insertion depth, (3) electrical dynamic range, and (4) signal processing strategy. There are also factors, such as patient's level of intelligence and communicativeness, which are unrelated to deafness that may also affect auditory performance. It is important to note, however, that much of the substantial intersubject variability that we observe in cochlear implant users remains unexplained.

Commercial Implant Processors.

There are currently three cochlear implant processors in the United States approved by the Food and Drug Administration (FDA), the Nucleus 24, the Clarion processor and the Med-El processor.

Nucleus 24 (CI24M).

The Nucleus 24 is the latest speech processor manufactured by Cochlear Pty. Limited, Australia. The processor can be programmed with either the SPEAK, the CIS or the Advanced Combined Encoder (ACE) strategies. In the SPEAK strategy, the signal is sent through a bank of 20 bandpass filters and out of the 20 filter outputs extracted through envelope detection, the 5-10 filter outputs with the largest amplitude are selected for stimulation. The selected electrodes (six on the average) are stimulated at a rate that varies from 100 to 300 pps. The CIS strategy does not use all 20 frequency bands available with the SPEAK strategy. A single frequency band is dedicated to each of the n electrodes, where n ranges from 4 to 12. The electrodes can be stimulated up to a maximum rate of 2400 pulses/sec depending on the value of n. This is a considerably higher rate than the rate (180-300 pulses/sec) used in the SPEAK strategy. The ACE strategy provides a faster (in terms of stimulation rate) implementation of the SPEAK strategy as the selected channels can be stimulated up to a maximum rate of 2400 pulses/sec.

The Nucleus 24 (CI24M) is available in two sizes, the regular size (SPrint) worn on the waist and the ear-level size worn behind the ear. The ear-level version (ESPrit) is the size of a behind-the-ear hearing aid and can be only programmed with the SPEAK strategy.

Clarion.

The Clarion cochlear implant system is the result of cooperative efforts among the University of California at San Francisco (UCSF), Research Triangle Institute (RTI) and the device manufacturer, Advanced Bionics Corporation (evolved from MiniMed Technologies). The Clarion implant supports a variety of speech processing options and stimulation patterns (Kessler, 1999). The stimulating waveform can be either analog or pulsatile. The stimulation can be simultaneous, sequential or both and the stimulation mode can be either monopolar or bipolar. The processor can be programmed with either the SAS strategy, the CIS strategy or the Paired Pulsatile Sampler (PPS) strategy. In the CIS strategy, pulses are delivered to 8 electrodes at a maximum rate of 833 pulses/sec per channel in an interleaved fashion. The PPS strategy is very similar to the CIS strategy, except that a pair of electrodes is stimulated simultaneously, thereby increasing the maximum rate per channel to 1666 pulses/sec. In the SAS strategy, the acoustic signal is processed through seven filters, compressed and then delivered simultaneously to seven electrode pairs. Pseudo-analog waveforms are delivered to each electrode at a rate of 13,000 samples/sec per channel. The Clarion processor also has the capability of supporting other speech processing strategies including a hybrid (CIS in some channels and SAS in others) strategy, however these strategies are still under investigation.

Med-El.

The Med-El cochlear implant processor is manufactured by Med-El Corporation in Austria. The Med-El processor (COMBI 40+) uses a 12-electrode array configured in monopolar mode and can be programmed with either the CIS or the n-of-m strategy. In the CIS strategy, the envelope amplitudes extracted from 12 bandpass filters are used to modulate biphasic pulses that are then delivered in an interleaved fashion to 12 monopolar electrodes at a maximum rate of 1,515 pulses/sec per channel.

The Med-El processor can also be programmed with a high-rate n-of-m strategy that is similar to the ACE strategy used in the CI24M device. In the n-of-m strategy, from a maximum of m channels, the n (n<m) channels with the highest energy are selected for stimulation in each cycle. Reducing the number of channels allows the stimulation rate per channel to increase, possibly leading to better temporal resolution of the sound signal without compromising spectral resolution. Due to limitations in RF transmission rates, it is sometimes necessary to reduce the number of electrodes to be stimulated in order to obtain higher rates of stimulation.

Clinical Results.

Adults. After the preceding description of cochlear implant systems and the signal processing and stimulation strategies they use, it may be of interest to review the kinds of perceptual benefit that patients typically achieve. Although many adult patients experience immediate results, scores in speech perception tests improve gradually over the first few weeks or months after implantation and then stay more or less constant. Figure 6 shows CNC word identification scores obtained by 29 postlingually deaf cochlear implant users grouped by the device they use. These results were obtained in the laboratory of the second author. All subjects had at least one year of experience with the device. Sets of 50 words were prerecorded and presented in open-set fashion at 70 dB SPL.

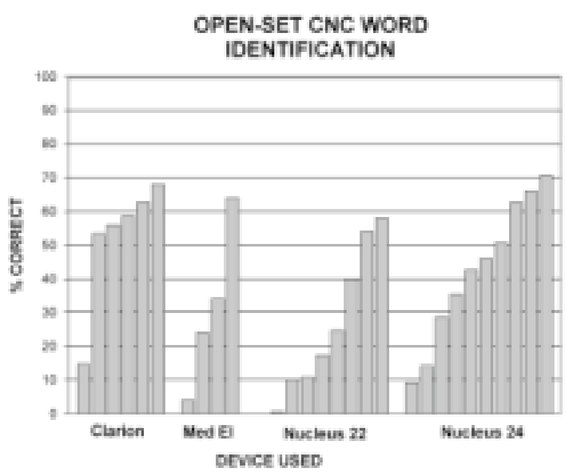


Figure 6. Perception of monosyllabic words without lipreading by postlingually, profoundly deaf adults with cochlear implants. These open-set scores were obtained in quiet, at least one year after initial activation of the patients' speech processors.

Subjects were not selected. Instead, they volunteered for studies that included several other tests. Due the nature of the study, patients using different devices were not matched along variables that may influence speech perception, such as duration of deafness, or amount of residual hearing. Therefore, it would be inappropriate to use these numbers to compare speech perception scores obtained with different devices. Instead, we can draw two conclusions that are valid for all four devices. First, almost all patients obtain some level of open set speech perception (only one out of 29 subjects was unable to identify any of the words). Second, there are large individual differences among patients. In this group of subjects, there were one Clarion user, one Med-El user, four Nucleus-22 users and two Nucleus-24 users who scored 15% words correct or less. On the other hand, there were several users of all devices who scored over 50% correct. A listener who can identify more than half of the words in a relatively difficult test like this that does not provide any linguistic context or allow for lipreading, probably can communicate in relatively fluent fashion in a face-to-face situation. Furthermore, the second author has observed many of these patients communicating fluently over the telephone, even when they didn't know who was calling or the nature of the call.

It is important to stress that the preceding results are representative of those obtained by postlingually deaf listeners. In contrast, speech perception scores are much worse for prelingually deaf adult cochlear implant users (people who become profoundly deaf before the age of three years, before their language skills are ingrained). Only rarely can a prelingually deafened adult cochlear implant recipient achieve any open-set speech perception. For example, Zwolan et al. (1994) report that none of the 17 subjects they tested were able to demonstrate any open-set speech recognition. Zwolan et al. (1996) also assessed changes in *closed-set* speech perception performance in 11 prelingually deaf adults before they received cochlear implants and one year after activation of the speech processor.

Seven tasks were used and tested for statistically significant differences at a 5% level using the binomial distribution. If changes pre- to post-implantation were completely random, it would be expected that about four comparisons (77 tests times the 5% significance level equals 3.85, which is rounded to 4) would yield "significant" increases and another four comparisons would yield "significant" decreases. Instead, Zwolan et al. found 11 significant decreases and four significant increases. Two of the increases were from 10% to 35% in a four-choice spondee test and from 12% to 30% in a four-choice final consonant test. In other words, the two post-increase scores may not be statistically different from the 25% that would be expected from chance alone. The Zwolan et al. article, which may be the largest most comprehensive and careful study of this population, found that most prelingually deafened patients who receive cochlear implants as adults used their devices regularly, were satisfied with them, and reported communicative benefit from its use. Nevertheless, it is clear that long-term sound deprivation that includes the first few years of life has a large, deleterious effect on speech perception after cochlear implantation. But what if a patient is congenitally or prelingually deaf and receives a cochlear implant within a few years after hearing loss, so that the period of sound deprivation is minimized?

Children. As we have indicated above, it has been well documented that implantation of congenitally or prelingually deaf children results in substantial benefit to speech perception (Meyer et al., 1998; Svirsky & Meyer, 1999; Meyer & Svirsky, 2000), and speech production (Svirsky et al., 2000b). Waltzman et al. (1997), in a study of 38 consecutively implanted, prelingually deafened and orally educated children showed that all subjects achieved significant open-set speech recognition. However, in the case of children who are born profoundly deaf, the auditory signal provided by a cochlear implant may also provide enough information for the development of oral language skills. In other words, one of the main potential outcomes of pediatric cochlear implantation (in addition to improvement of speech perception and speech production skills) is the development of skills in an oral language such as English. Numerous studies have found that prelingually deaf children lag in their English language abilities with respect to normal-hearing children.

A recent study (Svirsky et al., 2000a) assessed the expressive language of 23 prelingually deaf children before receiving cochlear implants as well as 6, 12 and 18 months after initial stimulation with a

Nucleus-22 device. Figure 7 shows the most important results of that study. The vertical axis indicates the children's expressive language skills, expressed in terms of age equivalent scores, and the horizontal axis shows their chronological age. Therefore, the linguistic development of an average child with normal hearing should follow the diagonal line. Black circles indicate the mean language age of the 23 patients at all four testing intervals and the white circles indicate the language age that would be expected at each interval if the children had not received cochlear implants. It is observed that at the preimplant interval, the children who would later receive cochlear implants were delayed with respect to children with normal hearing (because the first black circle is well below the diagonal). It may also be observed that these implanted children developed linguistic skills faster than would have been predicted for deaf children without implants (because the black circles fall above the white circles). Analysis of these and other data showed that language development over 2.5 years after implantation kept pace with the development that would have been expected from normal-hearing children at the same starting point of linguistic skills. In other words, the language gap that was present at implantation remained the same size (when measured in terms of language age) instead of increasing month after month as would be expected from deaf children without cochlear implants.

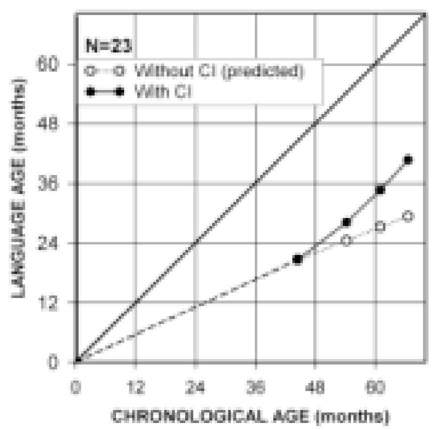


Figure 7. Expressive language skills (expressed in terms of language age) as a function of chronological age for 23 prelingually deaf cochlear implant users before implantation and at three intervals after implantation (black circles). White circles represent the expressive language growth expected from these children if they had not received cochlear implants. The solid diagonal line represents the scores expected from children with normal hearing.

Summary

Cochlear implants have been very successful in restoring partial hearing to profoundly deaf people. Most individuals with implants are now able to communicate and understand speech without lip-reading and many are able to talk over the telephone. The greatest benefits with cochlear implantation have occurred in patients who (1) acquired speech and language before their hearing loss and (2) have a shorter duration of deafness. Implant patients in United States can choose from three implant processors that support a variety of speech processing strategies. Clinicians now have the option to program the patient with multiple strategies and have the patient select the strategy they prefer. In addition, clinicians have the capability, thanks to the flexible fitting software, to change certain speech processing parameters to optimize performance.

References

Armstrong-Bendall G, Goodrum-Clarke K, Stollwerck L, Nunn T, Wei J, & Boyle P. (1999). Clarion paired pulsatile sampler (PPS): user preference and performance. *1999 Conference on Implantable Auditory Prostheses*, Pacific Grove, California.

Battmer R., Zilberman Y, Haake P, & Lenarz T. (1999). SAS-CIS pilot comparison study in Europe. *Annals of Otology, Rhinology & Laryngology* 108:69-73.

Brill S, Gstottner W, Helms J, Ilberg CV, Baumgartner W, Muller J, & Kiefer J. (1997). Optimization of channel number and stimulation rate for the fast CIS strategy in the COMBI 40+. *American Journal of Otology* 18: S104-6.

Blamey P, Arndt P, Bergeron F, Bredberg G, Brimacombe J, Facer G, Larky J, Lindstrom B, Nedzelski J, Peterson A, Shipp D, Staller S, & Whitford L. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology and Neuro-Otology*. 1:293-306.

Cohen NL, Roland JT, & Fishman A. (in press). Surgical technique for the Nucleus Contour cochlear implant. *Ear & Hearing.*

Hinojosa R, & Marion M. (1983). Histopathology of profound sensorineural deafness. *Annals of New York Academy of Sciences*. 405: 459-84.

Holden LK, Skinner MW, & Holden TA. (in press). Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear and Hearing*.

Kessler D. (1999). The Clarion multi-strategy cochlear implant. *Annals of Otology, Rhinology and Laryngology*. Suppl. 177. 108: 8-16.

Loizou P. (1998). Mimicking the human ear: an overview of signal processing techniques used for cochlear implants. *IEEE Signal Processing Magazine*. 15: 101-30.

Loizou P, Poroy O, & Dorman M.. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *Journal of the Acoustical Society of America*. 108: 790-802.

Merzenich M & White M. (1977). Cochlear implant - the interface problem. In: Hambrecht F and Reswick J, (eds.) *Functional electrical stimulation: applications in neural prostheses.* Pp. 321-340. Marcel Dekker.

Meyer TA, Svirsky MA, Kirk KI, & Miyamoto RT. (1998). Improvements in speech perception by children with profound prelingual hearing loss: effects of device, communication mode, and chronological age. *Journal of Speech, Language and Hearing Research*. 41: 846-58.

Meyer TA & Svirsky MA. (2000). Speech perception by children with the Clarion (CIs) or Nucleus 22 (SPEAK) cochlear implant or hearing aids. *Annals of Otology, Rhinology, & Laryngology.* Suppl.185: 49-51.

Nilsson M, Soli SD, & Sullivan J. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*. 95:1085-99.

Osberger M & Fisher L. (1999). SAS-CIS preference study in postlingually deafen adults implanted with the Clarion cochlear implant. *Annals of Otology, Rhinology & Laryngology.* Suppl. 177. 108: 74-79.

Parkinson AJ, Arcaroli J, Staller SJ, Arndt PL, Cosgriff A, & Ebinger K. (in press). The Nucleus 24 Contour cochlear implant system: adult clinical trial results. *Ear & Hearing.*

Plant K, Whitford L, Psarros C, Vandali A, & Clark G. (1999). Parameter selection and programming recommendations for the ACE and CIS speech processing strategies. *1999 Conference on Implantable Auditory Prosthesis*, Pacific Grove, CA.

Saunders E, Cohen L, Aschendorff A, Shapiro W, Knight M, Stecker M, Righter B, Waltzman S, Tykocinski M, Roland T, Laszig R, & Cowan R. (in press). Threshold, comfortable level and impedance changes as a function of electrode-modiolar distance. *Ear & Hearing.*

Svirsky MA, & Meyer A. (1999). A comparison of speech perception for pediatric Clarion and hearing aid users. *Annals of Otology, Rhinology, and Laryngology*. Suppl. 177. 108: 104-10.

Svirsky MA, Robbins AM, Kirk KI, Pisoni DB, & Miyamoto RT. (2000a). Language development in profoundly deaf children with cochlear implants. *Psychological Science*. 11: 153-58.

Svirsky MA, Sloan RB, Caldwell M, & Miyamoto RT. (2000b). Speech intelligibility of prelingually deaf children with multichannel cochlear implants. *Annals of Otology, Rhinology, & Laryngology*. Suppl. 185: 123-5.

van den Honert, & Stypulkowski P. (1987). Single fiber mapping of spatial excitation patterns in the electrically stimulated nerve. *Hearing Research*. 29: 195-206.

Vandali AE, Whitford LA, Plant KL, & Clark GM. (2000). Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear & Hearing*. 21:608-24.

Waltzman SB, Cohen NL, Gomolin RH, Green JE, Shapiro WH, Hoffman RA, & Roland JT, Jr. (1997). Open-set speech perception in congenitally deaf children using cochlear implants. *American Journal of Otology*. 18:342-9.

Wilson B. (1993). Signal processing. In: Tyler R, (ed.) *Cochlear Implants*: *Audiological Foundations*. Pp. 35-86. Singular Publishing Group, Inc.

Wilson B, Finley C, Lawson D, Wolford R, Eddington D, & Rabinowitz W. (1991). Better speech recognition with cochlear implants. *Nature*. 352: 236-38.

Wilson B, Lawson D, & Zerbi M. (1993). Speech Processors for auditory prostheses. NIH Project N01-DC-2-2401, *Fifth Quarterly Progress Report*.

Wilson B, Lawson D, & Zerbi M. (1995). Advances in coding strategies for cochlear implants. *Advances in Otolaryngology - Head and Neck Surgery*, 9: 105-29.

Ziese M, Stutzel A, von Specht H, Begall K, Freigang B, Sroka S, & Nopp P. (2000). Speech understanding with the CIS and N-of-M strategy in the MED-EL COMBI 40+ system. *ORL*. 62: 321-29.

Zwolan TA, Lougheed LA, & Kileny PR (1994). Clinical results of prelingually deafened adult cochlear implant patients. *Abstracts of the Seventeenth Midwinter Research Meeting, Association for Research in Otolaryngology*, February 6-10, 1994, St. Petersburg Beach, FL.

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

Effects of Short-Term Auditory Deprivation on the Control of Intraoral Air Pressure in Pediatric Cochlear Implant Users[1]

David L. Jones,[2] Mario A. Svirsky[3] and Sujuan Gao[4]

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

[1] This work was supported by NIH-NIDCD grant DC00423. We thank Allyson Riley for her assistance with data collection.
[2] University of Wyoming, Laramie, Wyoming.
[3] Indiana University School of Medicine and Purdue University, Indianapolis, IN.
[4] Indiana University School of Medicine, Indianapolis, IN.

Effects of Short-Term Auditory Deprivation on the Control of Intraoral Air Pressure in Pediatric Cochlear Implant Users

Abstract. The purpose of this study was to determine whether two speech measures –peak intraoral air pressure (IOP) and IOP duration – obtained during the production of intervocalic stops would be altered as a function of the presence (or absence) of auditory stimulation provided by a cochlear implant (CI). Five pediatric CI users were required to produce repetitions of the words *puppy* and *baby* with their CIs turned on. The implants were then turned off for 1 hour, at which time the speech sample was repeated with the implant still turned off. Seven normal-hearing (NH) children were used as a comparison group. They were also tested twice with a one-hour interval between sessions. Peak IOP and IOP duration were measured for the medial consonant in both conditions. The results show that auditory condition affected peak IOP more than IOP duration. Peak IOP was greater for /p/ than /b/ with the CI off, but some participants reduced or reversed this pattern when the CI was on. Our findings suggest that different speakers with CIs may employ different articulation strategies as they learn to use the auditory signal in speech production.

Introduction

The primary purpose of providing sound enhancement to an individual with a hearing impairment is to aid in the perception of sound and speech. At the same time, there is evidence to suggest that sensory devices such as hearing aids and CIs also serve to facilitate speech production. Speakers who have been provided with feedback from these devices have been shown to demonstrate shifts in various acoustic and perceptual variables, generally toward the range of normal (Kirk & Hill-Brown, 1985; Leder et al., 1986; Osberger, Maso, & Sam, 1993; Svirsky, Jones, Osberger, & Miyamoto, 1998; Tartter, Chute, & Hellman, 1989; Tobey & Hasenstab, 1991; Tye-Murray, Spencer, & Woodworth, 1995). This suggests that auditory feedback plays a critical role in improving speech intelligibility and maintaining acceptable speech in individuals with hearing impairment. Nonetheless, even in NH individuals, it is not clearly understood how auditory information is used to regulate specific articulatory, physiological, acoustic, or aerodynamic variables so that intelligible speech is developed and maintained.

The ability of adventitiously deafened adults to maintain reasonably intelligible speech has been presented as evidence that on-line auditory feedback is not essential once speech motor patterns have become well established (Borden, 1979; Zimmermann & Rettaliata, 1981). Conversely, the abnormal speech characteristics of children who are deafened early in life highlight the importance of auditory feedback in speech motor learning. A deaf infant or toddler is unable to monitor the acoustic consequences of her phonation. Without access to this source of feedback (Guenther, Hampson, & Johnson, 1998) it is very likely that perceptually critical speech parameters are either not regulated at all or regulated outside the range of activity that is required for perceptually acceptable speech. For a young pediatric CI user, the relationship between auditory feedback and speech motor learning is unique. Such a child is provided with auditory information only after maladaptive speech patterns have developed and needs to learn not only *how* to regulate specific speech parameters, but also *which* parameters to regulate so that perceptually acceptable speech may emerge.

In an effort to determine those speech production parameters that are invariant – and presumably most critical to the control of normal speech production – a number of studies have observed the response of the speech production system to various types of perturbation. Most of these studies have introduced some type of mechanical interference (see Folkins, 1985 for a review). Some studies have involved altering or interrupting feedback mechanisms that are believed to be operative for speech regulation (Kelso & Tuller, 1983; Lane & Tranel, 1971; Prosek & House, 1975; Siegel & Pick, 1974; Siegel, Pick, Olsen, & Sawin, 1976). Several studies have demonstrated changes in the speech production of NH individuals when auditory feedback is manipulated. When amplified speech is fed back to NH individuals, both adults (Siegel & Pick, 1974) and preschool children (Siegel et al., 1976) have been shown to decrease vocal intensity; whereas when auditory feedback is attenuated, vocal intensity is increased (Lane & Tranel, 1971). The immediacy of these responses to auditory feedback manipulations was interpreted as an indication that "an auditory feedback system is available and operative…but that it does not begin to regulate speech except under circumstances in which communication is difficult" (Siegel & Pick, 1974, p. 1624). However, the implications of much of this auditory feedback research tell us little about the role of auditory feedback in motor speech development because the individuals studied were already skilled speakers and the experimental paradigm used in these studies may have forced "attention to auditory feedback which may not necessarily operate under normal circumstances for skilled speakers" (Borden, 1979, p. 309).

"On-Off" Experiments

A more intriguing approach to auditory perturbation comes from a number of studies of induced short-term auditory deprivation in pediatric CI users. The removal of auditory feedback from a pediatric CI user likely reveals different phenomena than auditory perturbation in NH individuals. It may shed light on the manner in which the individual uses auditory feedback for speech production, the motor speech strategies that the individual has learned, and the degree to which the individual relies on online auditory feedback.

Tobey et al. (1991) tested 13 children using the on-off methodology and found that second formant frequencies of vowels were more centralized when the speech processor was turned off while it approached near-normal frequencies with the processor on. Richardson, Busby, Blamey, Dowell, and Clark (1993) studied vowel formant changes in three children and found no consistent changes across participants as a function of auditory feedback. Tye-Murray, Spencer, Bedia, and Woodworth (1996) used phonetic transcriptions and clinician's ratings to study the speech intelligibility of twenty pre-lingually deafened pediatric CI users in an on-off experiment. In general, the clinicians did not find any differences in intelligibility between the two auditory conditions. In some instances, however, the children exhibited vowel nasalization and inappropriate aspiration of initial consonants when the CI was on. The authors suggested that this reflected an attempt to increase proprioceptive feedback for the purpose of providing heightened awareness of their speaking behavior. Svirsky et al. (1998) used the on-off paradigm to study changes in nasalization in five pediatric CI users. In general, nasalization shifted toward the normal range with the CI turned on.

In two separtate studies of changes in phonation in two pediatric CI users, Higgins, McCleary, and Schulte (1999, 2001) found differences the children's responses to short-term (one hour) auditory deprivation. One child showed decreases in IOP for both /p/ and /b/, decreases in fundamental frequency and intensity, a decrease in nasal airflow for /m/, and an increased second formant for /a/ after the CI had been turned off. The other child showed no consistent effects for auditory deprivation with the exception of a reduced voice-onset-time for /p/. The authors noted that the child who showed the most consistent changes in the off condition demonstrated better speech production and perception skills, suggesting that this individual may have been using auditory feedback in a regulatory fashion.

Speech Aerodynamics and Auditory Feedback

Many of the articulatory abnormalities observed in the speech of the hearing-impaired have been attributed to aberrant control of speech aerodynamic parameters such as respiratory kinematics (Forner & Hixon, 1977) and oral airflow (Itoh, Horii, Daniloff, & Binnie, 1982; Whitehead & Barefoot, 1983). IOP, particularly in bilabial plosives, has also been a variable of interest in this population (Higgins, Carney, McCleary, & Rogers, 1996; Leeper, Perez, & Mencke, 1980). The production of bilabial plosives involves the rapid increase and release of impounded oral pressure. In normal speakers, the presence or absence of voicing differentiates the magnitude and timing of the IOP that is associated with the production of /b/ or /p/; the voiceless consonant /p/ is produced with higher peak IOP and longer IOP duration than its voiced counterpart /b/ (Arkebauer, Hixon, & Hardy, 1967; Bernthal & Beukelman, 1978; Subtelny, Worth, & Sakuda, 1966). Peak IOPs produced by hearing-impaired speakers have been found to be higher than normal (Martony & Lindqvist, as cited by Leeper et al., 1980), lower than normal (Leeper et al., 1980), or consistently negative (Higgins et al., 1996).

Although the control of speech aerodynamics may be important to developing and maintaining adequate speech, its specific role as a speech motor control variable is not clear. Vocal tract pressure has been proposed as a variable that is monitored and may be employed as part of a feedback system (Malecot, 1970; Prosek & House, 1975). Indeed, normal speakers are able to detect self-generated IOP differences as small as 1.01-1.12 cm H_2O (Williams, Brown, & Turner, 1987) – less than the IOP difference that distinguishes voiced and voiceless consonants. Warren (1986) hypothesized that regulation of vocal tract pressure is principal in speech motor control and that a speaker will maintain vocal tract aerodynamics even at the expense of acoustic accuracy. Moon and Folkins (1991) tested Warren's hypothesis by attempting to force NH speakers into a choice between aerodynamic and acoustic regulation. IOP was measured while the intensity level of frication was attenuated and fed back to the speaker online. Significant changes in IOP, IOP duration, and pressure curve area were observed, but only when the frication level was attenuated by at least 30 dB. In addition, the pressure changes were not as dramatic as the authors had predicted. Moon and Folkins concluded that their results did not support the notion of a regulating system that relies exclusively on speech aerodynamics, nor did it support a regulating system that relies exclusively on auditory feedback (as represented by level of frication).

The presence of a CI generally results in access to feedback not previously available; a deaf child must learn to use sensory input to modulate the speech control parameters that will lead to improved speech intelligibility. This introduction of brief periods of auditory deprivation in pediatric CI users may reveal the extent to which they have learned to modulate and maintain these parameters and the extent to which they may be using online auditory feedback to achieve auditory-perceptual goals. Although the regulation of speech aerodynamics is considered critical to the production of acceptable speech, there is relatively little information regarding the manner in which a child with a CI uses auditory feedback to control speech aerodynamic parameters such as IOP. The purpose of this study was to examine if short-term auditory deprivation alters the control of IOP in pediatric CI users. Specific questions of interest include:

1) When auditory information is provided, do peak IOPs and IOP durations approach the range of normal compared to when auditory information is unavailable?

2) When auditory information is available, are differences in the IOPs associated with the productions of /p/ and /b/ more distinct than when auditory information is unavailable (thus demonstrating the speakers' attempts to use auditory input to mark the /p/-/b/ contrast)?

3) When auditory information is unavailable, are the changes in IOP reflective of an increased reliance on tactile feedback?

Methods

Participants

The participants were five pediatric CI users and seven NH children who served as a comparison group. The CI users were selected for their ability to understand the instructions for this study and their good speech perception and production skills. The motivation for selecting children with superior speech skills (three of the five participants) was to maximize the potential for observing change in their speech when the speech processor was turned on and off. All of the CI user participants used the Nucleus-22 device. Table I shows demographic characteristics of these participants, including age at onset of profound deafness, age at implantation, age at which the speech production tests described here were administered, intelligibility and speech perception scores, and communication mode used by each participant. The participants had used their CIs for 1.5 to 3.7 years at the time of testing.

Table I. *Demographic information regarding the participants with CIs.*

Participant	CI1	CI2	CI3	CI4	CI5
Age at onset of deafness (years)	3.5	0.0	0.0	0.0	0.0
Age at implantation	7.3	8.9	5.5	6.1	5.3
Age at testing	9.9	10.8	9.2	7.6	7.5
Word intelligibility (Monsen Sentence Test)	91%	11%	71%	54%	74%
Speech perception (PBK-phoneme)	76%	74%	66%	84%	83%
Speech perception (LNT-easy words, phoneme)	82%	60%	81%	81%	95%
Communication mode	Oral	TC	TC	Oral	Oral

The seven NH participants all spoke Standard American English. They exhibited normal oropharyngeal anatomy, normal speech articulation, and normal nasal resonance as judged by both investigators. The age range of the normal participants was 8-11 years.

Background Information on the Participants with CIs. The CI users were part of a large group of implanted children whose speech perception and production skills are being followed longitudinally in our laboratory. As part of the longitudinal project, we obtain measures of their speech intelligibility, speech perception, and language development during periodic visits to our lab. The participants selected for this study agreed to undergo additional testing to assess their production of oral pressure during one of their regular sessions. The intelligibility scores and speech perception scores in Table I were obtained using methods described previously (Meyer & Svirsky, 2000; Meyer, Svirsky, Kirk, & Miyamoto, 1998; Svirsky et al., 1998; Svirsky, Sloan, Caldwell, & Miyamoto, 2000). Other studies performed with large, representative samples of pediatric CI users (e.g., Svirsky & Meyer, 1999) indicate that participants inthe present sample were more intelligible and understood speech better than the average pediatric CI user. Because of this, results from the present study may not be generalizable to the entire population of pediatric CI users.

Speech Sample

The speech stimuli included repetitions of the words “puppy” and “baby.” To minimize the articulatory complexity of the speech task for the CI users, a carrier phrase was not used for either group. The rate of repetition was modeled for all participants at approximately 1.5 syllables per second.

Instrumentation

IOP was measured with a rigid polyethylene (PE) tube that was placed between the participants’ lips and positioned in the medial aspect of the oral cavity. The tube was coupled to a pressure transducer (Microswitch 164PC01D37). The pressure signal was amplified and low-pass filtered at 50 Hz using a Biocommunication Electronics 201 amplifier, then digitized at a sampling rate of 2500 Hz. The pressure transducer was calibrated for centimeters of water (cm H_2O) pressure using a Glottal Enterprises MCU-2 manometer. The speech acoustic signal was transduced simultaneously using a Sony ECM-44B microphone and digitized. The acoustic signal was transduced solely for the purpose of token identification.

Procedure

The participants were provided with the speech sample using an oral model provided by the examiner as well as the signs for the target words ("puppy" and "baby"). The participants were instructed to practice the words prior to data collection to demonstrate that their utterances approximated the correct productions of the target words. The examiner did not correct or modify the participants' word productions because each participant was able to produce the target words with sufficient intelligibility.

Each CI user produced repetitions of each word while their CI was turned on (ON condition). The participants produced each word in sets of 5-6 repetitions until at least 25 tokens were obtained. The CI users were then instructed to turn off their speech processors and leave the testing area for one hour. At the end of one hour, the participants returned to the laboratory and the speech sample was repeated with the implant still turned off (OFF condition). The order of the conditions was consistent across participants and conditions, with repetitions of "puppy" being produced prior to repetitions of "baby." The sequence of the auditory feedback condition was always ON followed by OFF.

Tests on the NH participants were performed twice with a one-hour interval in between to simulate the timing of the ON and OFF conditions under which the CI users were tested. For the NH participants, the two conditions will also be referred to as ON and OFF with the understanding that NH participants were not subjected to any auditory deprivation. Normal participants practiced the speech sample prior to data collection in order to provide the "normal" range against which the CI data were compared.

Data Measurement and Analysis

Peak IOP and IOP duration for the medial consonants /p/ and /b/ were measured for each participant. Peak IOP was defined as the point at which the pressure signal achieved its maximum deviation from baseline. IOP duration was defined as the time between the onset and offset of IOP associated with production of the medial consonant. A token was considered to be acceptable if it had a well-defined single peak, and if its onset and offset both occurred relative to baseline. The first token in each set of repetitions was not measured. For each word production obtained under each auditory condition, it was our initial intent to measure twenty acceptable IOP tokens per participant; however, it was not possible to do so in every instance. When a token was determined to be unacceptable, it was due to a pressure offset secondary to saliva in the PE tube or a double-peaked signal, which is attributable to the tongue occluding the opening of the PE tube. We obtained the target number in 38 of 48 total instances in which measurement of 20 IOP tokens was targeted (12 total participants x 2 phonemic targets x 2 auditory conditions). In the remaining instances, the number of tokens measured was as follows: 6 instances of 19 acceptable tokens, 1 instance each of 18, 17, 15, and 14 acceptable tokens. The instances of 17, 15, and 14 measurable tokens occurred with three of the NH participants.

For each participant, mean peak IOPs and IOP durations were determined for both the ON and OFF conditions. General linear mixed models were used to conduct repeated measure analyses of variance with peak IOP and IOP duration as the dependent variables. Hearing status (CI vs. NH), auditory condition (ON-OFF), phonetic context (/p/ vs. /b/), and all interaction terms among the three factors were included as independent variables in the mixed models. A compound symmetry structure was assumed for the correlations for both the peak IOP and IOP duration models, where observations from the same child were treated as correlated and observations from different children were treated as independent. Post-hoc analyses (*t*-tests) were carried out to compare the mean peak IOP and IOP durations between the two auditory conditions and the two phonetic conditions for each participant. Given the total number of post-hoc comparisons made (48), a probability level of .001 was used for each post hoc comparison to control for multiple comparison errors and to achieve an overall significance level of .05 for all post-hoc tests.

Results

Mean IOP Duration

Results of the mixed model analysis for mean IOP duration are provided in Table II. Table III shows the mean IOP durations obtained for the group of CI users in the ON and OFF conditions and for the group of NH speakers in the ON and OFF conditions. A significant three-way interaction between hearing status, condition and phonetic context was obtained. Post-hoc analyses showed that for productions of /p/, the CI group exhibited higher mean IOP durations than the NH group in both the ON and OFF conditions. However, only the OFF condition reaching statistical significance ($t(223)=7.79$, $p=.0001$). Within the CI group, mean IOP duration was longer in the OFF condition than in the ON condition; however, this difference was not statistically significant. For productions of /b/, the mean IOP durations for the CI group were longer than the mean for the NH group in both the ON and OFF conditions; the differences were statistically significant for both auditory conditions ($t(235)=4.47$, $p=.0001$ for ON; $t(236)=6.47$, $p=.0001$ for OFF). Within the CI group, no significant difference in IOP duration was found between the ON and OFF conditions.

Table II. *Results of general linear mixed model analysis.*

Variable: IOP Duration

Effect	Num DF	Den DF	F value	Pr > F
STATUS (CI--NH)	1	10	2.99	0.1143
CONDITION (ON--OFF)	1	10	0.13	0.7227
CONTEXT (/p/--/b/)	1	10	790.69	<.0001
STATUS*CONDITION	1	10	15.76	0.0026
STATUS*CONTEXT	1	10	30.96	0.0002
CONDITION*CONTEXT	1	10	2.39	0.1533
STATUS*CONDITION*CONTEXT	1	10	14.37	0.0035

Variable: Peak IOP

Effect	Num DF	Den DF	F value	Pr > F
STATUS (CI--NH)	1	10	9.62	0.0112
CONDITION (ON--OFF)	1	10	29.32	0.0003
CONTEXT (/p/--/b/)	1	10	258.26	<.0001
STATUS*CONDITION	1	10	2.17	0.1713
STATUS*CONTEXT	1	10	23.48	0.0007
CONDITION*CONTEXT	1	10	10.20	0.0096
STATUS*CONDITION*CONTEXT	1	10	7.70	0.0196

The results of post-hoc testing of the individual effects of auditory condition and phonetic context are presented below.

Auditory Condition: ON versus OFF. The mean IOP durations obtained for each of the NH and CI user participants in the ON and OFF conditions are shown in Table III. During productions of /p/, three participants (CI1, CI4, and CI5) showed no significant difference between the ON and OFF conditions. Two participants (CI2 and CI3) showed a statistically significant difference, with mean IOP duration being longer in the OFF condition for both participants (69.6 ms longer for CI2, 29.15 ms longer for CI3). Two of the NH participants (NH1 and NH3) showed statistically significant ON-OFF differences in mean IOP duration, with IOP duration being longer during ON (47.75 ms difference for NH1, 51.15 ms difference for NH3).

During productions of /b/, four participants (CI1, CI2, CI3 and CI4) showed no significant difference between auditory conditions. One participant (CI5) exhibited significantly longer IOP durations in the OFF condition (23.55 ms difference). Two of the NH participants showed statistically significant ON-OFF differences in mean IOP duration, with participant NH6 showing a longer mean IOP duration during the ON condition (25.6 ms difference), and participant NH2 showing a longer mean IOP duration in the OFF condition (18.05 difference).

Table III. *Individual and Group* mean *IOP durations (in ms) and standard deviations for productions of /p/ and /b/ by NH and CI user participants. Same-letter superscripts represent post hoc comparisons that reached statistical significance. Lowercase superscript: p≤ .001; Uppercase superscript: p≤.0001.*

	/p/		/b/	
	ON	OFF	ON	OFF
CI1	192.15 (16.74) [A]	184.60 (18.46) [B]	129.35 (15.44) [A]	127.20 (14.88) [B]
CI2	194.05 (42.67) [C]	263.65 (35.54) [C,D]	157.28 (25.30)	138.45 (11.83) [D]
CI3	133.65 (9.22) [E]	162.80 (18.78) [E,F]	122.75 (17.15)	129.20 (18.66) [F]
CI4	323.47 (53.52) [G]	337.79 (63.11) [H]	167.79 (41.80) [G]	137.32 (33.33) [H]
CI5	192.25 (28.41) [J]	187.65 (21.18) [K]	133.45 (14.42) [I,J]	157.00 (16.50) [I,K]
	ON	OFF	ON	OFF
NH1	177.95 (20.10) [L,M]	130.20 (13.56) [L,N]	112.35 (14.43) [M]	105.25 (8.46) [N]
NH2	164.10 (15.32) [P]	170.65 (14.89) [Q]	118.85 (13.39) [o,P]	136.90 (17.74) [o,Q]
NH3	208.55 (34.07) [R,S]	157.40 (15.44) [R,T]	120.95 (17.48) [S]	120.30 (14.79) [T]
NH4	147.50 (11.28) [U]	141.85 (11.58) [V]	117.05 (13.93) [U]	114.65 (10.47) [V]
NH5	175.68 (68.12)	191.23 (57.28) [W]	144.10 (22.60)	130.70 (13.12) [W]
NH6	230.65 (35.42) [Y]	217.64 (57.52) [Z]	152.65 (19.53) [X,Y]	127.05 (10.88) [X,Z]
NH7	173.50 (21.54) [AA]	176.40 (20.77) [BB]	119.80 (15.76) [AA]	116.50 (13.74) [BB]

Group Means

	/p/		/b/	
	ON	OFF	ON	OFF
CI	205.94 (70.41)	226.18 (73.31) [CC]	141.55 (29.81) [DD]	137.84 (22.54) [EE]
NH	182.61 (42.30)	166.24 (40.53) [CC]	126.54 (21.93) [DD]	121.58 (16.23) [EE]

Phonemic target: /p/-/b/ Contrast. In both the ON and OFF conditions, all participants exhibited IOP durations that were longer for /p/ than /b/; the difference was statistically significant for every comparison across all participants except participants N5, CI2, and CI3 in the ON condition. For the CI users, the magnitude of the /p/-/b/ difference within participants was similar across auditory condition for all participants except CI2, where the /p/-/b/ difference was much greater in the OFF condition than it was in the ON condition. In addition, it is evident that participant CI4 exhibited mean IOP durations for /p/ that were dramatically longer than those for /b/ in both auditory conditions.

Mean Peak IOP

Results of the mixed model analysis for peak IOP are provided in Table II. Table IV shows a summary of the mean peak IOPs obtained for the CI and NH groups in the ON and OFF conditions. A significant three-way interaction between hearing status, condition and phonetic context is again shown. For productions of /p/*,* the CI users produced mean peak IOPs that were significantly higher than the mean peak IOPs for the NH group for both the ON and OFF conditions ($t(236)=9.59$, $p=.0001$ for ON; $t(225)=9.37$, $p=.0001$ for OFF). Within the CI group, the mean peak IOP obtained in the OFF condition was higher than the mean peak IOP in the ON condition, but this difference was not statistically significant. For productions of /b/, the mean peak IOPs in both the ON and OFF were higher for the CI group than the NH group. The differences were statistically significant in both auditory conditions ($t(235)=15.38$, $p=.0001$ for ON; $t(234)=73.36$, $p=.0001$ for OFF). No significant difference in mean peak IOP was shown for /b/ within the CI group between the ON and OFF conditions.

Table IV. *Individual and Group mean peak IOPs (in cmH_2O) and standard deviations for productions of /p/ and /b/ by the CI and NH participants. Same-letter superscripts represent post hoc comparisons that reached statistical significance. Lowercase superscript: $p \leq .001$; Uppercase superscript: $p \leq .0001$.*

	/p/		/b/	
	ON	OFF	ON	OFF
CI1	9.70 (1.39) [A,B]	14.08 (2.15) [A]	11.70 (1.48) [B]	12.48 (1.06)
CI2	16.24 (3.43) [c]	20.08 (2.85) [c,D]	18.92 (2.48)	16.22 (2.84) [D]
CI3	8.40 (1.32)	7.33 (1.97)	7.81 (1.17) [E]	5.69 (1.10) [E]
CI4	8.14 (1.64) [f]	9.95 (1.34) [f]	7.52 (1.03)	8.48 (2.02)
CI5	16.31 (4.39) [G]	14.02 (2.04) [H]	9.48 (1.94) [G]	11.08 (1.24) [H]
	ON	OFF	ON	OFF
NH1	7.11 (0.98) [I]	8.50 (0.75) [I,J]	5.96 (1.34)	4.77 (1.66) [J]
NH2	9.32 (1.24) [k,L]	11.49 (1.69) [k,M]	5.40 (1.18) [L]	6.72 (1.83) [M]
NH3	11.08 (0.70) [O]	10.85 (2.21) [P]	4.89 (1.85) [N,O]	7.06 (1.74) [N,P]
NH4	7.02 (0.71) [Q]	6.72 (0.64) [R]	3.58 (1.34) [Q]	4.45 (1.19) [R]
NH5	5.72 (1.72)	6.21 (1.03) [S]	4.02 (1.29)	4.27 (1.08) [S]
NH6	6.30 (1.40) [u]	6.74 (1.07)	4.81 (1.08) [u,T]	6.71 (0.74) [T]
NH7	6.00 (0.70) [V,W]	8.48 (0.85) [V,X]	4.63 (1.04) [W]	5.51 (1.74) [X]

Group Means

	/p/		/b/	
	ON	OFF	ON	OFF
CI	11.79 (4.61) [AA]	13.12 (4.84) [BB]	10.96 (4.43) [CC]	10.82 (4.00) [DD]
NH	7.52 (2.14) [AA]	8.55 (2.34) [BB]	4.76 (1.49) [CC]	5.64 (1.82) [DD]

The results of the post-hoc tests of group differences and the individual effects of auditory condition and phonetic context are presented below.

Auditory Condition: ON versus OFF. In general, the within-target comparisons reveal that some CI users demonstrated higher mean IOPs in the ON condition, while others showed higher IOPs in the OFF condition. For productions of /p/, three of the five CI users (CI1, CI2, and CI4) exhibited mean peak IOPs that were significantly higher in the OFF condition than in the ON condition (by magnitudes ranging from 1.81-4.38 cmH_2O). Participants CI3 and CI5 demonstrated higher mean IOPs in the ON condition, although neither of the comparisons for these participants showed a statistically significant difference. Some of the NH participants exhibited statistically significant ON-OFF differences, with participants NH1, NH2, and NH7 showing higher mean peak IOPs in the OFF condition (by magnitudes ranging from 1.39-2.48 cmH_2O).

For productions of /b/, two CI users (CI2 and CI3) showed higher mean IOPs in the ON condition. Only CI3 showed a statistically significant difference (2.12 cmH_2O difference). Participants CI1, CI4, and CI5 showed higher mean pressures in the OFF condition; none of these differences was statistically significant. For the NH participants, mean peak IOPs were significantly higher in the OFF condition for participants NH3 and NH6 (by magnitudes of 2.17 and 1.90 cmH_2O, respectively).

Phonetic Target: /p/-/b/ Contrast. For the NH participants, mean peak IOPs were higher during productions of /p/ than /b/; the differences were statistically significant in all instances except one (NH6, OFF condition). For the CI users, comparison of mean peak IOPs for /p/ and /b/ within auditory condition are presented in Figure 1. The mean peak IOP was consistently higher for /p/ than for /b/ in all participants in the OFF condition (by magnitudes ranging from 1.47–3.86 cmH_2O); the difference was statistically significant only for participants CI2 and CI5. This effect was not observed in the ON condition, however, where only one participant (CI5) had a mean peak IOP that was significantly higher for /p/ than /b/ (by 6.83 cmH_2O). For participants CI3 and CI4, mean peak IOP for /p/ was slightly higher than it was for /b/ (by 0.59 and 0.62 cmH_2O, respectively), but these differences were reduced compared to the OFF condition, and were not statistically significant. The remaining participants (CI1 and CI2) exhibited mean peak IOPs for /b/ that were higher than those for /p/ (by 2.00 and 2.68 cmH_2O, respectively); this difference was statistically significant for CI1. For these two participants, a reversal in the peak IOP contrast occurred for /p/ and /b/ in the OFF condition.

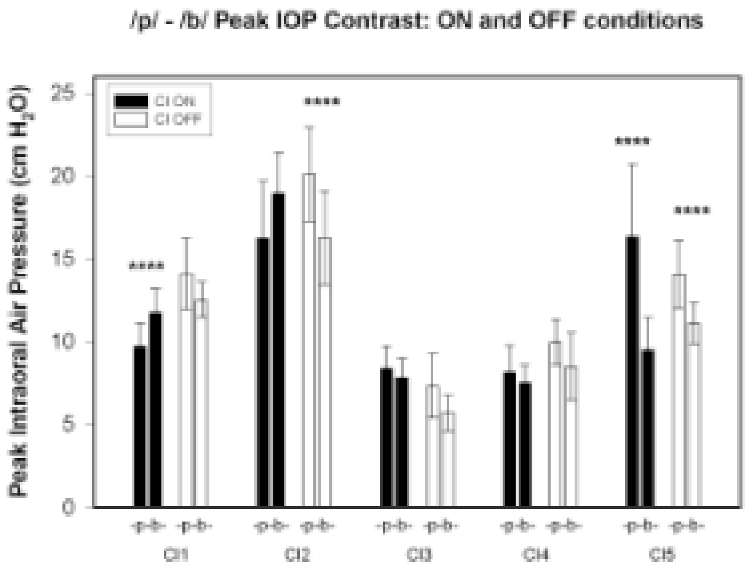


Figure 1. Comparison of mean peak IOPs for /p/ versus /b/ within auditory condition for the CI users. Statistically significant post-hoc comparisons are marked with asterisks. (**** $p \leq .0001$).

Discussion

The purpose of this study was to determine if two specific aerodynamic variables in speech production would be altered by auditory feedback provided by a CI. Given the exploratory nature of this study, explicit hypotheses regarding how the variables of interest might be altered were not proposed. Nonetheless, we anticipated that the variables would shift toward the range of normal when auditory information was available (i.e., when the CI was on). Such a shift did not occur on a consistent basis. In the ON condition, the CI participants exhibited mean peak IOPs for /p/ that were lower and closer to normal than in the OFF condition; however, there was essentially no ON-OFF difference for productions of /b/. All participants (both CI and NH) showed appropriately higher peak IOPs for /p/ than for /b/ in the OFF condition, but this contrast was reduced or reversed in the ON condition for four of the five CI users. This counterintuitive but interesting result has no precedent in the CI literature.

Of the four participants (CI1 thru CI4) who reduced or lost the /p/-/b/ contrast for peak IOP in the ON condition, no two participants exhibited this outcome in the same way. Relative to the OFF condition, these participants showed the following IOP patterns with the CI turned on: participant CI1 maintained a stable IOP for /b/, but decreased IOP for /p/ considerably; participant CI2 decreased the IOP for /p/ considerably while increasing the IOP for /b/ significantly; participant CI3 increased IOP for both /p/ and /b/, but the magnitude of the increase was greater for /b/; and participant CI4 decreased IOP for both /p/ and /b/, but the magnitude of the decrease was greater for /p/. Some of these inter-subject differences may reflect the participants' efforts to shift the magnitude of their peak IOPs to within normal limits, but this is not always an adequate explanation. For instance, Participant CI2's mean IOP shifted in the direction of normal for productions of /p/, but it shifted away from normal for /b/. In addition, the mean peak IOPs for participants CI3 and CI4 were already within the range of normal, so no significant IOP manipulation was required.

IOP duration did not appear to vary as a function of the presence or absence of auditory feedback. While a statistically significant difference was observed between the ON and OFF conditions, the magnitude of the difference was usually not very large (30 ms or less). Further, although the mean IOP duration values were higher in the CI group, the magnitudes of the CI participants' ON-OFF differences were comparable to the ON-OFF differences exhibited by the NH participants. Thus, just as in the NH participants, it is possible that the few ON-OFF differences observed in the CI participants were due to variability that was related to subtle changes in word duration.

The ON-OFF manipulation did not influence the /p/-/b/ contrast for IOP duration in the same way that it did for peak IOP. Mean IOP duration was greater for /p/ than it was for /b/ across both auditory conditions for all participants; this contrast was statistically significant in 13 of the 14 comparisons within the NH group, and in 8 of the 10 comparisons within the CI group. It is possible that the participants may have exploited IOP duration rather than peak IOP in an attempt to mark the voiced-voiceless distinction. The difference in mean peak IOP between /p/ and /b/ has been well established in NH children and adults, and the implication exists that this peak IOP difference is integral to producing and perceiving the voiced-voiceless distinction. However, it has been suggested that both the longer IOP duration and the higher peak IOP that occur during production of the voiceless /p/ help to passively suppress vocal fold vibration for a longer period relative to /b/ (Boucher & Lamontagne, 2001; Stevens, 1991), resulting in a longer voice onset time.

If this is true, then the CI participants who reversed the /p/-/b/ contrast for peak IOP may have employed one of two strategies. It is possible that they may have tried to maintain relatively stable IOP distinctions in duration to influence voice onset time, thereby marking the /p/-/b/ contrast. The results of this study lend tentative support to this notion. A second strategy that participants CI1 and CI2 may have employed to override the higher peak IOPs produced for the voiced consonant /b/ is the implementation of an active laryngeal adjustment to initiate voicing to counter the higher peak IOPs they produced. This type of strategy implies that some children with CIs might modify their speech production by selectively exploiting the acoustic dimensions of the speech signal that are well conveyed by the implant. Participants CI1 and CI2 may have exhibited relatively higher subglottal pressures for the medial /b/ in the word "baby" (where all segments are voiced) secondary to an attempt to increase the intensity of the voicing parameter, which is well perceived by implant users. The auditory information provided from the alternating voiceless-voiced-voiceless-voiced sequence in "puppy" is less consistent; attempts to accentuate the silence associated with the voiceless consonants may have led to a disproportionate decrease in SPL (Svirsky, Lane, Perkell, & Wozniak, 1992), resulting in lower subglottal pressures for /p/ relative to /b/. Thus, the IOP differences observed in the ON condition may be an indirect consequence of the speakers' attempts to manipulate a parameter that is associated with speech aerodynamics, such as SPL (Stathopoulos, 1986), and not the speakers' attempts to actively control an aerodynamic parameter, such as peak IOP.

Interestingly, all participants demonstrated an appropriate /p/-/b/ distinction regarding peak IOPs in the OFF condition. This finding suggests that the participants may have a greater reliance on tactile feedback with the CI off. Given that all of the participants were implanted after the age of 5, this may reflect a control strategy that is more fully developed, allowing for more consistent speech production at this stage in their implant use. Even though these individuals had from 1.5 to 3.7 years of experience with their CIs, it is likely that these participants were still in the process of learning to use the auditory signal to match specific articulatory and aerodynamic processes with a desired acoustic-perceptual signal. In essence, they may be in the process of replacing a well-established speech production model (one without auditory feedback) with a new modified model (one with auditory feedback).

The statistically significant ON-OFF differences that were observed in some of the NH participants were not anticipated. Despite these observed differences, it should be noted that when the mean peak IOPs and IOP durations of the NH group were collapsed across conditions, the mean values fall within the range of mean values reported for normal children (Bernthal & Beukelman, 1978; Leeper et al., 1980; Subtelny et al., 1966). Further, regardless of the ON-OFF differences exhibited by the individuals within this group, appropriate /p/-/b/ contrasts were maintained consistently across conditions. These unexpected within-subject differences could reflect either the fact that NH children tend to demonstrate greater variability in IOP production than do adults (Bernthal & Beukelman, 1978; Subtelny et al., 1966) or a practice effect, as familiarity with the speech task may have resulted in an alteration of speech behaviors. The likelihood of these results occurring in the NH participants may also have been increased by the fact that we did not control for sound pressure level during the speech task, which was motivated by concern that it would lead to speech production patterns by the CI users that would have been uncharacteristic of them in the OFF condition. Although a practice effect may have influenced the results obtained for the CI users, this is less likely since the two conditions were different for this group. Certainly, randomizing the order of the auditory condition across the CI users would have helped to address this issue. However, scheduling of the CI users' regular clinical evaluations on the same day did not permit us to manipulate the auditory condition in this way.

Implications

The small number of participants studied and the relatively inconsistent nature of the results precludes any firm generalizations regarding how pediatric CI users make use of auditory feedback in the regulation of speech production. There appeared to be no relationship between the results reported above and duration of CI use. Participant CI3 had the most experience and exhibited IOP values that were within normal limits, but did not demonstrate the appropriate /p/-/b/ contrast for peak IOP. Of the two individuals who completely reversed the peak IOP contrast with the CI on, one (CI1) had the best word intelligibility and the other (CI2) had the worst word intelligibility among the CI users. The only participant who exhibited the appropriate /p/-/b/ contrast for both IOP measures was CI5; this participant also demonstrated the best speech perception skills in this group. Keeping in mind the somewhat heterogeneous makeup of the CI group studied, the results of this study, as well as others (e.g., Higgins et al., 1999, 2001), are most likely representative of the fact that regardless of age of implantation or experience with the device, children with CIs may be found at different stages of learning what information to extract from the auditory signal for the purpose of developing and regulating speech production.

It is not clear how a developing child uses auditory input to develop perceptually acceptable speech. Guenther et al. (1998) posited that speakers use an auditory perceptual reference frame to plan speech movements for vowels and semivowels. Callan, Kent, Guenther, and Vorperian (2000) suggest that auditory feedback of self-produced speech may be critical to developing speech movements that reproduce auditory targets. Although it is not the intent of this study to validate the auditory perceptual model of speech motor control, this model provides some insight into the speech production of children with CIs. Our assumption is that children with implants develop a model of speech production (i.e., a model of the acoustic and linguistic consequences of their articulatory actions) that is unlike the model developed by NH children. The two models may be similar for some of the acoustic dimensions that are easy to perceive by CI users, such as temporal cues (Van Tasell, Greenfield, Logemann, & Nelson, 1992) and changes in the speech waveform envelope (Van Tasell, Soli, Kirby, & Widin, 1987). However, the models may be different from normal in two possible ways. First, the CI model may not include those acoustic dimensions that are more difficult for CI users to perceive, such as rapidly changing spectral cues (Dorman, Dankowski, McCandless, Parkin, & Smith, 1991). Second, the CI model may include acoustic dimensions that are easily perceived by CI users, but which are overly relied upon for marking articulatory distinctions and which do not necessarily improve speech production (e.g., enhancing the voice-voiceless distinction by increasing the low-amplitude voicing signal for /b/ relative to /p/).

The results of this study are of interest in regard to Warren's (1986) theory that vocal tract pressure serves as the basic parameter of speech motor control. If this were the case, then one would have expected to see no change in peak IOPs as a function of auditory feedback. This pattern would be expected for children with profound hearing loss. These children are considered to rely heavily on tactile feedback during speech development. One might expect that the range of vocal tract pressures that are learned early in the speech development process would become ingrained and resistant to manipulation – even when auditory stimulation is provided. Alternatively, it could be assumed that the CI users in this study had access to an auditory signal for a duration sufficient to learn to regulate peak IOP within a relatively stable range. The fact that the peak IOP was altered as a function of auditory feedback implies that this aerodynamic parameter may not play a primary role in speech motor control, at least in children who were deafened prelingually or at a young age. At the same time, the relative stability of IOP duration that was observed across auditory conditions suggests that temporal aspects of vocal tract pressure may be a more critical control parameter, even if it is not regulated within an operating range that is close to normal (e.g., participant CI4's IOP durations for /p/). It should be noted that Warren, Morr, Rochet, and Dalston (1989) recognized the importance of auditory feedback to the development of speech production and suggested that vocal tract pressure becomes a fundamental control parameter *after* its association with auditory feedback is established during the process of speech development in the presence of normal hearing.

Two methodological issues deserve attention with regard to this study. First, although differences were observed in the aerodynamic variables as a function of the presence or absence of auditory information, there is some question as to whether a longer period of auditory deprivation would have resulted in greater changes in the measures of interest. Continued manipulation of this variable will be a topic of future investigation. Second, we did not attempt to correlate the aerodynamic differences to changes in speech intelligibility. However, as Tye-Murray et al. (1996) discuss, children with prolonged CI use exhibit changes in speech production using the ON-OFF paradigm that are not necessarily detectable to a listener. Our participants had an average of approximately 24 months experience with their CIs, which is not drastically less than the average use of 34 months by the participants in the Tye-Murray et al. study. Although perceptual speech ratings were not obtained as part of this study, changes in speech intelligibility and voicing errors were not noted by the primary investigator (DLJ) in the OFF condition. It is possible that a longer period of auditory deprivation would bring about greater aerodynamic changes with co-occurring changes in speech intelligibility. At the same time, it is acknowledged that parameters other than IOP are involved in the production of an intelligible /p/ or /b/. Further studies will be necessary to address these issues in speech production.

The results of the present study are also relevant to the role of auditory feedback in ongoing speech production. It has been proposed that auditory input is crucial for the development of ongoing speech production during the early stages of learning, but that continuous auditory feedback becomes less important as mature speech motor patterns develop (Guenther, et al., 1998; Tobey, 1993; Zimmermann & Rettaliata, 1981). We hypothesize that speech production in implanted children will be strongly dependent on the auditory information provided by the implant at early stages after implantation. Assuming that a pediatric CI user has the ability to extract acoustic information for the purpose of regulating speech production, this dependence on auditory feedback should decrease over time as users acquire more stable motor patterns for speech production. The ON-OFF experimental paradigm allows a unique opportunity to test this hypothesis. We would expect to see large differences between the ON and

OFF conditions shortly after implantation; these ON-OFF differences should decrease with time as the speech motor patterns become more embedded.

References

Arkebauer, H.J., Hixon, T.J., & Hardy, J.C. (1967). Peak IOPs during speech. *Journal of Speech and Hearing Research*, *10*, 196-208.

Bernthal, J.E. & Beukelman, D.R. (1978). IOP during the production of /p/ and /b/ by children, youths, and adults. *Journal of Speech and Hearing Research*, *21*, 361-371.

Borden, G.J. (1979). An interpretation of research on feedback interruption in speech. *Brain and Language*, *7*, 307-319.

Boucher, V. & Lamontagne, M. (2001). Effects of speaking rate on the control of vocal fold vibration: Clinical implications of active and passive aspects of devoicing. *Journal of Speech, Language, and Hearing Research*, *44*, 1005-1014.

Callan, D.E., Kent, R.D., Guenther, F.H., & Vorperian, H.K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research*, *43*, 721-736.

Dorman, M.F., Dankowski, K., McCandless, G., Parkin, J.L., & Smith, L. (1991). Vowel and consonant recognition with the aid of a multichannel CI. *The Quarterly Journal of Experimental Psychology*, *43A*, 585-601.

Folkins, J.W. (1985). Issues in speech motor control and their relation to the speech of individuals with cleft palate. *Cleft Palate Journal*, *22*, 106-122.

Forner, L.L. & Hixon, T.J. (1977). Respiratory kinematics in profoundly hearing-impaired speakers. *Journal of Speech and Hearing Research, 20*, 373-408.

Guenther, F.H., Hampson, M. & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, *105*, 611-633.

Higgins, M.B., Carney, A.E., McCleary, E., & Rogers, S. (1996). Negative IOPs of deaf children with CIs: physiology, phonology, and treatment. *Journal of Speech and Hearing Research*, 39, 957-967.

Higgins, M.B., McCleary, E.A., & Schulte, L. (1999). Altered phonatory physiology with short-term deactivation of children's CIs. *Ear & Hearing*, *20*, 426-438.

Higgins, M.B., McCleary, E.A., & Schulte, L. (2001). Articulatory changes with short-term auditory deactivation of the CIs of two prelingually deafened children. *Ear & Hearing*, *22*, 29-45.

Itoh, M., Horii, Y., Daniloff, R.G., & Binnie, C.A. (1982). Selected aerodynamic characteristics of deaf individuals during various speech and nonspeech tasks. *Folia Phoniatrica*, *34*, 191-209.

Kelso, J.A. & Tuller, B. (1983). "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research*, *26*, 217-224.

Kirk, K.I. & Hill-Brown, C. (1985). Speech and language results in children with a CI. *Ear and Hearing*, *6*, 36S-47S.

Lane, H. & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, *14*, 677-709.

Leder, S.B., Spitzer, J.B., Milner, P., Flevaris-Phillips, C., Richardson, F., & Kirchner, J.C. (1986). Reacquisition of contrastive stress in adventitiously deaf speakers using a single channel CI. *Journal of the Acoustical Society of America*, *79*, 1967-1974.

Leeper, H.A., Perez, D.M., & Mencke, O. (1980). The influence of utterance length upon bilabial closure duration of selected deaf children. *Journal of Communication Disorders*, *13*, 373-383.

Malecot, A. (1970). The lenis-fortis opposition: Its physiological parameters. *Journal of the Acoustical Society of America*, *47*, 1588-1592.

Meyer, T.A. & Svirsky, M.A. (2000). Speech perception by children with the Clarion (CIS) or Nucleus (SPEAK) CI or hearing aids. *Annals of Otology, Rhinology, and Laryngology*, *109* (Supp 185), 49-50.

Meyer, T.A., Svirsky, M.A., Kirk, K.I., & Miyamoto, R.T. (1998). Improvements in speech perception by children with profound prelingual hearing loss: Effects of device, communication mode, and chronological age. *Journal of Speech, Language, and Hearing Research*, *41*, 846-858.

Moon, J.B. & Folkins, J.W. (1991). The effects of auditory feedback on the regulation of IOP during speech. *Journal of the Acoustical Society of America*, *90*, 2992-2999.

Osberger, M.J., Maso, M., & Sam, L. (1993). Speech intelligibility of children with CIs, tactile aids, or hearing aids. *Journal of Speech and Hearing Research*, *36*, 186-203.

Prosek, R.A. & House, A.S. (1975). IOP as a feedback cue in consonant production. *Journal of Speech and Hearing Research*, *18*, 133-147.

Richardson, L.M., Busby, P.A., Blamey, P.J., Dowell, R.C., & Clark, G.M. (1993). The effects of auditory feedback from the Nucleus CI on the vowel formant frequencies produced by children and adults. *Ear and Hearing*, *14*, 339-349.

Siegel, G.M. & Pick, H.L. (1974). Auditory feedback in the regulation of voice. *Journal of the Acoustical Society of America*, 56, 1618-1624.

Siegel, G.M., Pick, H.L., Olsen, M.G., & Sawin, L. (1976). Auditory feedback in the regulation of vocal intensity of preschool children. *Developmental Psychology*, *12*, 255-261.

Stathopoulos, E.T. (1986). Relationship between IOP and vocal intensity in children and adults. *Journal of Speech and Hearing Research*, *29*, 71-74.

Stevens, K.N. (1991). Vocal-fold vibration for obstruent consonants. In J.Gauffin & B. Hammarberg (Eds.), *Vocal-fold physiology: Acoustic, perceptual, and physiological aspects of voice mechanisms* (pp.29-36). San Diego, CA: Singular.

Subtelny, J.D., Worth, J.H., & Sakuda, M. (1966). IOP and rate of flow during speech. *Journal of Speech and Hearing Research*, *9*, 498-518.

Svirsky, M.A., Jones, D., Osberger, M.J., & Miyamoto, R.T. (1998). The effect of auditory feedback on the control of oral-nasal resonance by pediatric CI users. *Ear & Hearing*, *19*, 385-393.

Svirsky, M.A., Lane, H., Perkell, J.S., & Wozniak, J. (1992). Effects of short-term auditory deprivation on vowel formant frequencies in multichannel CI users. *Journal of the Acoustical Society of America*, *89*, 2895-2904.

Svirsky, M.A. and Meyer, T.A. (1999). A comparison of speech perception for pediatric Clarion and hearing aid users. *Annals of Otology, Rhinology, and Laryngology*, *108* (Sup. 177), 104-110.

Svirsky, M.A., Sloan, R.B., Caldwell, M., and Miyamoto, R.T. (2000). Speech intelligibility of prelingually deaf children with multichannel CIs. *Annals of Otology, Rhinology, and Laryngology*, *109* (Sup. 185), 123-125.

Tartter, V.C., Chute, P.M., & Hellman, S.A. (1989). The speech of a postlingually deafened teenager during the first year of use of a multichannel CI. *Journal of the Acoustical Society of America*, *86*, 2113-2121.

Tobey, E.A. (1993). Speech production. In R.S. Tyler (Ed.), *CIs: Audiological Foundations* (pp. 257-316). San Diego, CA: Singular Publishing.

Tobey, E.A., Angelette, S., Murchison, C., Nicosia, J., Sprague, S., Staller, S.J., et al. (1991). Speech production performance in children with multichannel CIs. *American Journal of Otology*, *12*, 165-173.

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

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

Tye-Murray, N., Spencer, L., & Woodworth, G. (1995). Acquisition of intelligible speech by children who have prolonged CI experience. *Journal of Speech and Hearing Research*, *38*, 327-337.

Van Tasell, D.J., Greenfield, D.G., Logemann, J.J., & Nelson, D.A. (1992). Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of CIs. *Journal of the Acoustical Society of America*, *92*, 1247-1257.

Van Tasell, D.J., Soli, S.D., Kirby, V.M., & Widin, G.P. (1987). Speech waveform envelope cues for consonant recognition. *Journal of the Acoustical Society of America*, *82*, 1152-1161.

Warren, D.W. (1986). Compensatory speech behaviors in individuals with cleft palate: A regulation/control phenomenon? *Cleft Palate Journal, 23*, 251-260.
Warren, D.W., Morr, K.E., Rochet, A.P., & Dalston, R.M. (1989). Respiratory response to a decrease in velopharyngeal resistance. *Journal of the Acoustical Society of America*, *86*, 917-924.
Whitehead, R.L. & Barefoot, S.M. (1983). Airflow characteristics of fricative consonants produced by normally hearing and hearing-impaired speakers. *Journal of Speech and Hearing Research*, *26*, 185-194.
Williams, W.N., Brown, W.S., & Turner, G.E. (1987). IOP discrimination by normal-speaking subjects. *Folia Phoniatrica*, *39*, 196-203.
Zimmermann, G. & Rettaliata, P. (1981). Articulatory patterns of an adventitiously deaf speaker: Implications for the role of auditory information in speech production. *Journal of Speech and Hearing Research*, *24*, 169-178.

The Indiana Speech Project: An Overview of the Development of a Multi-Talker Multi-Dialect Speech Corpus[1]

Cynthia G. Clopper, Allyson K. Carter, Caitlin M. Dillon, Luis R. Hernandez, David B. Pisoni, Connie M. Clarke,[2] James D. Harnsberger,[3] and Rebecca Herman[4]

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

[1] This work was supported by NIH-NIDCD Research Grant DC00111 to Indiana University.
[2] University of Arizona, Tucson, AZ.
[3] University of Florida, Gainesville, FL.
[4] DePaul University, Chicago, IL.

The Indiana Speech Project: An Overview of the Development of a Multi-Talker Multi-Dialect Speech Corpus

Abstract. The goal of the Indiana Speech Project (ISP) was to collect a corpus of spoken language samples from a number of different talkers that represent several different regional varieties of American English. Audio recordings were made of five college-aged women from each of six geographic regions of Indiana while they read isolated words, sentences, and a passage and while engaged in a conversation with an experimenter. The residential histories of the women and those of their parents were strictly controlled to ensure that the talkers were good representatives of each dialect region. The Indiana speech corpus will be used for acoustic-phonetic measurements of speech and perceptual studies on regional language variation by different groups of listeners.

Introduction and Theoretical Motivation

In recent years, researchers in the field of speech perception and spoken language processing have developed a number of speech corpora for conducting acoustic and perceptual experiments on human language. These corpora typically included a number of speakers reading a set of words or sentences [e.g., "Easy-Hard" Word Multi-Talker Speech Database (Torretta, 1995); TIMIT Acoustic-Phonetic Continuous Speech Corpus (Zue, Seneff, & Glass, 1990); Talker Variability Sentence Database (Karl & Pisoni, 1994)]. Although the talkers often included both males and females, other important indexical variables such as socioeconomic status, age, ethnicity, and regional dialect were rarely, if ever, considered in selecting the talkers. One exception to this rule is the TIMIT Acoustic-Phonetic Continuous Speech Corpus. This corpus contains spoken sentence materials from 630 talkers, representing eight different regional dialects of American English (Zue et al., 1990). While the TIMIT database was originally collected for speech recognition research, it has been used in various acoustic-phonetic studies on the role of gender, age, and dialect in linguistic variation (e.g., Byrd, 1992; Byrd, 1994; Keating, Blankenship, Byrd, Flemming, & Todaka, 1992; Keating, Byrd, Flemming, & Todaka, 1994).

While this work was going on in speech perception and speech recognition, sociolinguists have been collecting speech samples from talkers of a variety of ages, socioeconomic statuses, ethnicities, and regional dialects. The emphasis in this research has been on capturing the variability in spoken language as well as collecting extensive demographic information on each talker [e.g., Santa Barbara Corpus of Spoken American English (DuBois, Chafe, Meyer, & Thompson, 2000); CallFriend Telephone Speech Corpus for American English (Linguistic Data Consortium, 1996); TELSUR (Labov, Ash, & Boberg, in press)]. In contrast to the speech stimuli used in speech perception and speech recognition research, these speech samples are typically taken from "natural" language situations such as interviews and telephone calls and less emphasis is placed on obtaining identical utterances from the same set of talkers. While these corpora are useful for many kinds of sociolinguistic research, they are not adequate for perceptual research in which consistent linguistic content across talkers is highly desirable.

The initial goal of the Indiana Speech Project (ISP) was to collect a large amount of speech from a number of phonologically distinct dialect regions in the state of Indiana for use in perceptual studies and acoustic analyses. We wanted to combine the best aspects of the speech perception corpora with the unique focus of the sociolinguistic corpora. In particular, our goal was to collect a large corpus of utterances that were consistent across all talkers, allowing for better control of the stimulus materials for a wide range of perceptual and acoustic studies. In addition, we also wanted the talkers in our corpus to be homogeneous on all variables except regional dialect, which we manipulated following the strict guidelines described in detail below.

The TIMIT corpus was originally collected for speech recognition research, which means that it lacks two important elements for speech perception research. First, while each talker in the TIMIT corpus was recorded reading ten sentences, only two test sentences were spoken by all 630 talkers. As a result, researchers interested in comparing talkers without introducing variation in linguistic content are limited to just two sentences. Second, each talker in the TIMIT corpus was assigned to one of eight regional dialect labels, but the precise details of how those assignments were made is not explained in the accompanying documentation. In general, the dialect labels were assigned based on where the talker spent the majority of his or her childhood (William Fisher, personal communication), but there was little, if any, specific information collected about the regional dialect of the talker's parents. In addition, the labels used to define the dialects of the TIMIT talkers and the map used to define the regions (see Fisher, Doddington, & Goudie-Marshall, 1986) does not correspond with any current sociolinguistic theory of American English regional dialect variation.

Our speech corpus was designed to be an improvement over the TIMIT corpus because it was designed specifically for perceptual and acoustic research on dialect variation in American English. First, the ISP corpus was designed such that every talker was recorded reading the same materials with the exception of a short spontaneous speech sample. Second, the talkers in the ISP corpus were carefully screened on a number of variables related to their own residential history, as well as that of their parents, in order to ensure that the dialect labels they were assigned were appropriate. Thus, the two major features of this new corpus are the controlled nature of the stimulus set and the procedures used to carefully screen, select, and assign the talkers to dialect categories.

The ISP corpus will allow us to continue and extend the research on the perception of dialect variation that we have been conducting using the TIMIT corpus (e.g., Clopper, 2000; Clopper & Pisoni, 2002). In particular, this new corpus will allow us to further investigate dialect identification and categorization, dialect discrimination by both native and non-native speakers, the role of dialect variation in lexical decision and word recognition studies, dialect intelligibility, and the role of dialect variation in perceptual learning of novel voices and novel dialects. In addition, the controlled nature of the stimulus items we have recorded will allow us to conduct more detailed acoustic analyses of the vowel systems of different regional varieties, including the study of diphthongs, and to consider some of the acoustic correlates of stress as a function of dialect.

Talkers

Thirty adult women were recruited from the Bloomington campuses of Indiana University and Ivy Tech State College to serve as talkers for the ISP corpus. All of the women were monolingual, native speakers of American English with no history of hearing or speech disorders reported at the time of recording. The women ranged in age from 18 to 22 years old, with a mean age of approximately 20. Both the mother and the father of each participant were also native speakers of American English. Twenty-nine of the 30 talkers were Caucasian (non-Hispanic) and the remaining talker was Hispanic in ethnicity but was not fluent in Spanish. All of the talkers also had limited experience with foreign languages (typically restricted to classroom instruction). They also had little, if any, formal knowledge of linguistics or linguistic research methods. All participants were paid $25 ($10/hour) and were given a free Speech Research Laboratory t-shirt for their time.

Female talkers were selected for use in this corpus for several reasons. First, women are much more cooperative as participants in studies of this kind that require several sessions of an hour to an hour and a half in length. Second, sociolinguists have found that women tend to be ahead of men in language changes in progress, regardless of whether the changes are above or below the level of conscious social awareness (Labov, 1990; Milroy & Milroy, 1993). Speech stimuli recorded from female talkers might therefore be expected to reveal current changes in progress, such as the Northern Cities shift in Chicago or the Southern vowel shift in Louisville.

The talkers were specifically recruited to represent six different dialect regions of Indiana: Fort Wayne, Indianapolis, Bloomington, Evansville, and the Indiana counties near Chicago, IL and Louisville, KY. The map in Figure 1 shows the state of Indiana and its 92 counties. The regions were defined as the county containing the city of interest and all contiguous counties. In the case of Louisville, the region was defined as those counties in Indiana that are contiguous with the county in Kentucky in which Louisville is located. In the case of Chicago, the region was defined as the three counties in Indiana that, like Chicago, are on Central Time. No participants were actually from either Chicago or Louisville. In Figure 1, the county in which the city of interest is located is shown in dark gray and the regions are shown in light gray. All participants were required to have lived exclusively in the region they were representing prior to attending school in Bloomington. In order to participate, both of the talker's parents had to have grown up in a county contiguous with or within the region of interest. For example, a participant from Evansville (located in Vanderburgh County) had to live exclusively in Posey, Vanderburgh, Warrick, and/or Gibson counties and both of her parents had to have grown up in any of those counties and/or Spencer, DuBois, Pike, or Knox counties.

Regional varieties of the English spoken in Indiana were selected for the ISP for several reasons. First, given our location in Bloomington, Indiana at Indiana University, we had access to a large number of undergraduate students from all over the state of Indiana. Second, Indiana has historically been more diverse linguistically and culturally than its neighbors on either side, Illinois and Ohio. In particular, an examination of the settlement patterns in the early 19th century reveals that roughly the southern half of Indiana was first settled by people migrating from Kentucky and points farther south who followed the Ohio River and its tributaries into southern Indiana (Bergquist, 1981). The linguistic result of this settlement pattern is the extension of Southern English dialect features in both phonology and the lexicon into much of southern Indiana (Carmony, 1965; Gibbens, 1962). Because such Southern English features are not found in southern Illinois or southern Ohio, the northward shift of the major north-south dialect boundary in Indiana is called the "Hoosier Apex" (Carver, 1987).

We therefore expected to find sizable differences in pronunciation patterns between northern and southern Indiana speakers as a result of the major north-south boundary that runs through Indiana. This major dialect boundary should separate Chicago, Fort Wayne, and Indianapolis from Bloomington, Evansville, and Louisville. In addition, Chicago has long been identified as one of the cities in which the Northern Cities vowel shift is taking place (Labov, Yeager, & Steiner, 1972), so we expected the Chicago suburbs in Indiana to reflect this change as well. At the other extreme, we predicted that the most southern features would be found in the speech of the women from the Louisville suburbs, given their proximity to Kentucky. We expected the remaining cities to fall on a continuum of variation between these two endpoints, reflecting a gradual shift from northern to southern varieties of American English (Davis & Houck, 1992).

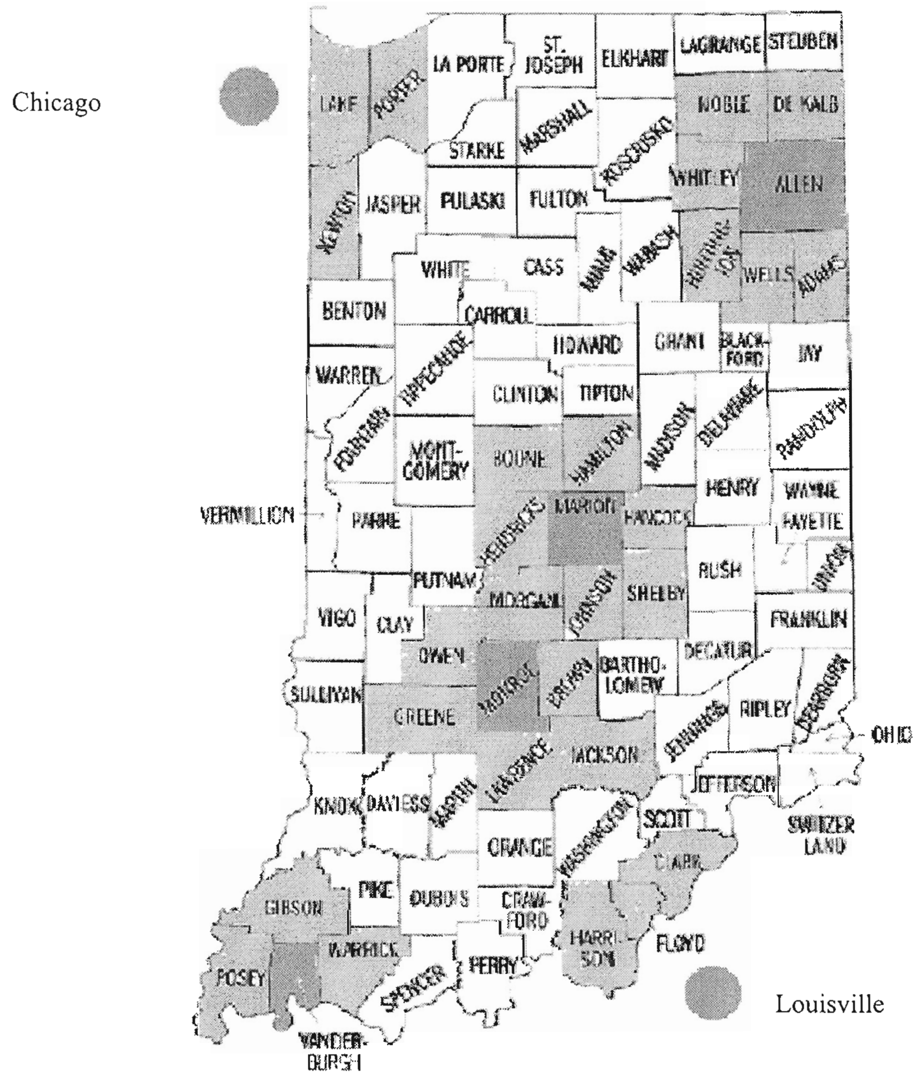


Figure 1. Map of the state of Indiana, including county lines. The dark counties represent those containing a city of interest: Allen (Fort Wayne), Marion (Indianapolis), Monroe (Bloomington), and Vanderburgh (Evansville). The dark circles outside the state line represent the other two cities of interest: Chicago, IL (to the northwest) and Louisville, KY (to the southeast). The light gray counties are those that belong to the region surrounding one of the six key cities.

Materials

Each talker was recorded while she read eight different sets of test materials. Recordings were also made of a spontaneous speech sample obtained from a short conversation with the experimenter. These materials are shown in Table 1 and are described in more detail in the sections below.

Materials Set	Number of Tokens	Examples
CVC Words	991	mice, dome, bait
Multisyllabic Words	240	alfalfa, nectarine
Disyllabic Nonwords	56	deploze, soggith
High Probability SPIN Sentences	200	All the flowers were in bloom. For your birthday I baked a cake.
Low Probability SPIN Sentences	100	Ruth will consider the herd. David does not discuss the hug.
Anomalous Sentences	100	Bill knew a can of maple beads. We're super so let's fish in the map.
Vowel Space	10 (10 repetitions)	heed, hid, head
Rainbow Passage	1	
Spontaneous Speech	1 (5 minutes)	

Table 1. Spoken language materials recorded by each talker for the ISP.

CVC Words. In order to study the acoustic properties of segmental variation in as many potentially relevant linguistic contexts as possible, it was necessary to collect a large number of consonant-vowel-consonant (CVC) utterances. The CVC list was composed of 915 different CVC words selected from an online dictionary containing approximately 20,000 entries based on Webster's Pocket Dictionary (Luce & Pisoni, 1998). The list is composed of all of the CVC words in the dictionary that received an average familiarity rating of 6.0 or greater (on a 7-point scale) by undergraduates (Nusbaum, Pisoni, & Davis, 1984). From the full set of these CVC words, 76 were selected to be produced twice, for a total of 991 tokens in the CVC list. The "repeated CVC" subset was selected such that every monophthong in English occurred at least five times in the list and every diphthong occurred at least four times in the list. Additionally, all diphthongs that participate in a shift or merger in one of the major regional dialects of American English occurred at least five times (see Labov et al., in press and Thomas, 2001 for discussions of vocalic variants and mergers in American English). All of the vowels with the exception of /oɪ/ occurred at least five times in this repeated CVC list. The final consonants in these words were selected such that contexts of dialect interest for each vowel were represented at least once. Contexts of interest for a given vowel were based on expected dialectal variation due to documented shifts and mergers (see Labov et al., in press and Thomas, 2001 for documentation of vocalic variation in American English). Following consonants were also varied systematically to include voiceless stops, voiced stops, nasals, liquids, and sometimes fricatives. The CVC words that were produced twice in the CVC list are shown in Appendix 1.

Multisyllabic Words. In order to study the effects of segmental interactions due to consonant clusters and stress, we also included a short list of multisyllabic words in the corpus. The multisyllabic word list consisted of a subset of the test items developed by Carter and Clopper (2000) for studying word reduction behavior in adult populations. The list was balanced for number of syllables, primary stress location, and morphological complexity. Table 2 shows the distribution of the words in this list with respect to these three variables. The list contained eight sets of 30 words of two-, three-, and four-syllables with primary stress equally distributed on the first, second, and third syllable. Half of the words in each set were monomorphemic and half were polymorphemic. All of the words were randomly selected from an online dictionary containing approximately 20,000 entries based on Webster's Pocket Dictionary. In addition, all of the words received a familiarity rating of at least 6.0 (on a 7-point scale) from undergraduate students, had a lexical frequency rating of one standard deviation from the log mean frequency, and had a neighborhood density of 2 or lower (Nusbaum et al., 1984). These criteria mean that the multisyllabic words were highly familiar, commonly occuring English words with few phonologically similar words that they might be confused with.

Number of Syllables	Primary Stress Location	Morphological Complexity	Number of Words
2	1st	monomorphemic	15
		polymorphemic	15
	2nd	monomorphemic	15
		polymorphemic	15
3	1st	monomorphemic	15
		polymorphemic	15
	2nd	monomorphemic	15
		polymorphemic	15
	3rd	monomorphemic	15
		polymorphemic	15
4	1st	monomorphemic	15
		polymorphemic	15
	2nd	monomorphemic	15
		polymorphemic	15
	3rd	monomorphemic	15
		polymorphemic	15

Table 2. Distribution of words in the multisyllabic word list with respect to number of syllables, primary stress location, and morphological complexity.

Disyllabic Nonwords. A set of disyllabic nonwords was also included in the corpus in order to compare American English vowels in the stressed syllable position of words with two different stress patterns, initial stress (Strong-weak or Sw) and final stress (weak-Strong or wS). The nonwords consisted of 56 disyllables representing fourteen American English vowels, /i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ʌ/, /a/, /u/, /ʊ/, /oʊ/, /ɔ/, /ai/, /aʊ/, and /oɪ/. All of these vowels except /oʊ/, /ai/, /aʊ/, and /oɪ/ were placed in those conditioning contexts in which particular dialect variants or mergers occur (see Labov et al., in press and Thomas, 2001 for discussions of vocalic variants and mergers in American English).

The 56 nonwords were designed to elicit the two different stress patterns (initial or final stress) by manipulating their syllabic structure. Twenty-eight of the nonwords (two tokens of each vowel) were designed to elicit initial stress ("Sw" words) by taking the form of a CVCVC sequence in which the medial C was always ambisyllabic. Such a "balanced" sequence of consonants and vowels is predicted to elicit initial stress because of stress biases in the American English lexicon. For example, in a study by Cutler and Carter (1987) of a corpus of 190,000 English words, 90% of content words were found to begin with "strong" syllables (e.g., clóset, óctopus, and sálamànder), which correspond to initial stress in our disyllabic nonwords.

The other half of the nonword stimulus set consisted of 28 disyllabic nonwords (two tokens of each vowel) designed to elicit final stress ("wS" words). Given the bias in English for disyllabic words to bear initial stress, it was critical to make use of certain phonological and syntactic patterns in English in order to elicit a consistent final stress pronunciation. First, CVCCVC sequences were used for all nonwords with predicted final stress. The medial consonant clusters in these sequences were always permissible word onset clusters in English, to encourage listeners to parse the nonwords as a CV.CCVC. Given the role that syllable weight normally plays in stress assignment in English and other languages, this parsing should elicit second syllable stress for these words (Hammond, 1999; Hayes, 1995).

Finally, for all nonwords, two additional criteria were adopted to ensure that the stimuli were sufficiently different from any real English words and that they would be produced with the intended vowel qualities and stress assignments. First, neither syllable in any of the disyllabic nonwords constituted a real English word orthographically (including proper names, but not slang). Second, all syllables involved phonotactically permissible segment sequences in English.

To encourage the appropriate stress assignment in production, the Sw and wS nonwords were placed in sentential contexts that were predicted to favor either initial or final stress, respectively. It has been shown experimentally that grammatical position affects stress assignment in nonwords (Kelly, 1988; Kelly & Bock, 1988). In one experiment, participants were presented with disyllabic nonsense words in various sentential contexts. The results showed that the talkers were more likely to pronounce nonsense words in noun contexts with a Sw pattern (Kelly and Bock, 1988). In another experiment, participants were presented with tape-recorded disyllabic nonwords and asked to construct sentences with these items (Kelly, 1988). Participants were significantly more prone to place words with final stress (wS) in verbal positions. Given these results, the nonwords designed to elicit initial stress were placed in a noun context, namely the subject position in the carrier sentence, "The ____ chased the ball." The nonwords designed to elicit final stress were placed in a verb context in the carrier sentence, "He will _____ the cookie dough."

Development of the nonword stimuli required two pilot experiments, using a total of 23 participants. In the first pilot, 13 participants were asked to produce the first iteration of the 56 nonwords in the carrier sentences discussed above. The 56 trials were randomized across initial stress and final stress. Analysis of stress location accuracy revealed that 40 of the 56 forms were produced with a stress accuracy rate of 86% or more, while 16 were produced with a stress accuracy rate of 78% or less. We then conducted a second pilot experiment in which we changed the orthography of the 16 words that were below the accuracy criterion of 86% to avoid problematic segmental effects that seemed to be common to all 16 below-criterion forms. We also changed the procedure such that trials were presented in two blocks. Subjects were first presented with all nonwords with predicted final stress, followed by all nonwords with predicted initial stress. This ordering was chosen to discourage talkers from carrying over the preferred initial stress pattern from noun context sentences to verb context sentences (with predicted final stress nonwords). Within blocks, individual trials were randomized. Ten participants were presented with all 40 of the "old" nonwords that had an accuracy rate over 86%, as well as the new, altered set of 16 nonwords. Results showed that each of the "new" nonwords reached an accuracy rate above the 86% criterion. The disyllabic nonword list consisted of these 56 nonwords from the two pilots and is shown in Appendix 2. Examples of the nonwords are shown in Table 1.

High and Low Probability SPIN Sentences. In addition to isolated words, we also included three sets of sentences for use in acoustic and perceptual studies of variation. Two of the three sentence lists were taken from the Speech Perception in Noise (SPIN) test (Kalikow, Stevens, & Elliot, 1977). These sentences are 5 to 8 words in length and are phonetically balanced based on phoneme frequency in English. The final word in each of the sentences is the target word. There are eight SPIN lists composed of 50 sentences each. Twenty-five of the sentences in each list are termed high probability (HP) because the target word is predictable from the semantic context of the sentence. The other 25 sentences in each list are termed low probability (LP) because the target word in each of these sentences is not predictable from the meaning of the sentence. All 200 of the high probability SPIN sentences were recorded in the ISP corpus. Half of the low probability sentences were chosen for the ISP. In particular, all 25 low probability sentences from lists 1, 2, 7, and 8 of the original SPIN test were selected, for a total of 100 low probability SPIN sentences. Examples of the SPIN sentences are shown in Table 1.

Anomalous Sentences. One hundred semantically anomalous sentences were created based on the SPIN sentences (Kalikow, Stevens, & Elliot, 1977). Target words for high and low probability sentences in the SPIN test are matched across lists such that high probability target words in list 1 match low probability target words in 2 and vice versa. The same relationship holds for lists 7 and 8. Therefore, in order to create anomalous sentences with the same target words as the low probability sentences that had already been selected for the ISP, the sentence structure and target words were taken from the 25 high probability sentences in the original SPIN lists 1, 2, 7, and 8. All content words (nouns, verbs, adjectives, adverbs, and some prepositions) of the original sentences were replaced with randomly selected content words that were the same part of speech from high probability sentences in lists 3, 4, 5, and 6. The resulting 100 sentences were semantically anomalous, but syntactically correct.

In order to confirm that the anomalous sentences could be read by untrained talkers without difficulty or disfluency, the anomalous sentences were presented visually to 13 participants who were recorded reading them aloud. Seven sentences were read by more than three of the talkers with disfluencies and had to be revised. In order to confirm that the anomalous sentences were all roughly equivalent in terms of their semantic anomaly, the revised list was then presented visually to 10 participants who rated each sentence on a 7-point sensibleness/strangeness scale. The five sentences which fell outside of one standard deviation from the mean rating were revised. This second revised list was then presented visually to 11 new participants who again rated them on their sensibleness/strangeness on a 7-point scale. All of the revised sentences fell within a single standard deviation of the mean rating. This third list was presented visually to ten participants who read them aloud, to ensure that revisions had not caused any of the sentences to be more difficult to read aloud fluently. One of these revised sentences was read with disfluencies by more than three talkers and was revised once again. The final anomalous sentence list used in the ISP corpus was thus the result of several successive stages of revision from the original list due to both perception and production pilot studies. After this process was completed, the list consisted of 100 sentences with the same target words as the 100 low probability SPIN sentences. The anomalous sentences were also all roughly equal with regard to their semantic anomaly and could be produced by naïve talkers fluently. Examples of the anomalous sentences are shown in Table 1.

Vowel Space. The vowel space portion of the corpus consisted of a set of familiar English words that can serve as benchmarks for mapping out the vowel system of a given individual talker or dialect group. This data will allow us to directly compare the vowel spaces of talkers from the different regions of Indiana, while keeping linguistic context constant across all of the utterances. Vowel spaces are typically mapped in two-dimensions, corresponding to first and second vowel formant frequencies (Gerstman, 1968; Peterson & Barney, 1952). The stimulus materials for this portion of the corpus consisted of 10 monosyllabic words, *heed*, *hid*, *aid*, *head*, *had*, *hut*, *odd*, *who'd*, *hood*, and *owed*, representing ten American English vowels, /i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ʌ/, /a/, /u/, /ʊ/, and /oʊ/, respectively. The use of these particular test words was motivated by the relatively constant consonantal context in which the vowels of interest appear (hVd or _Vd, where *V* = vowel). Each of the ten items was repeated ten times during the vowel space block of the experiment.

Passage of Connected Speech. A short continuous speech sample was obtained from each talker as a controlled, read passage which would allow for detailed analysis of individual variability, as well as dialectal differences at a higher level of language use than isolated words or sentences. Several passages were considered, including the Arthur the Rat passage (used to elicit dialectal differences for the Dictionary of American Regional English; Cassidy, 1985) and the Grandfather Passage (used by researchers of speech motor disorders; cf. Darley, Aronson, & Brown, 1975). However, the passage selected for the current project was the initial paragraph of the Rainbow Passage shown in Appendix 3 (Fairbanks, 1940). The Rainbow Passage has been used in numerous acoustic and perceptual studies since its first publication, including investigations into speaker differences (e.g., Gelfer & Schofield, 2000), individual variability (e.g., Sapienza, Walton, & Murry, 1999), and clinical populations (e.g., Hillenbrand & Houde, 1996; McHenry, 1999; Baker, Ramig, Johnson, & Freed, 1997).

Spontaneous Speech. In addition to recording subjects reading words, sentences, and the passage, we also collected a five-minute sample of spontaneous speech from each participant while she was engaged in a conversation with the experimenter in order to provide another somewhat more "natural" comparison to the read speech collected with the other materials. The spontaneous speech sample was based on the participants' responses to questions concerning their demographic background, such as information about their hometown, family, hobbies, and interests.

Methods

Participants were recorded individually in two separate test sessions that took place on different days. The first session lasted approximately an hour and a half, including breaks. The second session lasted approximately an hour, including breaks. Each set of materials was presented as a single experimental block. In all seven of the tasks in which there was more than one trial (i.e., all but the passage and the spontaneous speech tasks), the individual items were presented in a different random order for each participant. Prior to beginning a new block of trials, the participants were given instructions and were encouraged to ask questions if they did not understand what they were expected to do. On the first day of recording, talkers first read the CVC words, followed by the vowel space, multisyllabic words, and spontaneous speech tasks. The latter three tasks were performed in random order by each participant. On the second day, the talkers read the remaining materials (all three sentence lists, the disyllabic nonwords, and the passage). The order of the tasks on the second day was varied randomly across talkers. Breaks were given within and between each experimental block, as needed.

During all of the recording blocks, the participants were seated in a sound-attenuated chamber (IAC Audiometric Testing Room, Model 402) in front of a ViewSonic LCD monitor (ViewPanel VG151) mirroring the screen of a Macintosh Powerbook G3 which the experimenter, who was also seated in the sound-attenuated chamber, held on his or her lap. The participants read the materials off of the monitor screen as they were presented, speaking into a head-mounted Shure microphone (SM10A) positioned at the left corner of the mouth less than an inch from the face. The microphone signal was amplified by a tube microphone preamplifier (Applied Research Technology) which was connected to a Roland UA-30 USB Audio Interface which digitized the amplified microphone signal. The output of the Roland UA-30 was sent to the Powerbook and to Audio-Technica headphones (ATH-M2X) worn by the experimenter so that he or she was able to hear the signal as it was being sent to the Powerbook.

For the real words and sentence materials, a single word or sentence was displayed on the screen on each trial and the participant was asked to read the item aloud. The amount of time that the participant had to respond to the test item on the screen varied for the different materials. These recording times are shown in Table 3. All recordings were digitized at a sampling rate of 44.1kHz and each item was recorded into an individual .aiff sound file on the Powerbook. If the participant made an error in reading the item or if the experimenter heard any extraneous noise through the headphones, the experimenter pressed the "r" key on the Powerbook keyboard and the trial was repeated again at the end of the block.

Materials Set	Recording Time
CVC Words	2250 ms
Multisyllabic Words	3500 ms
Disyllabic Nonwords	5000 ms
High Probability SPIN Sentences	5000 ms
Low Probability SPIN Sentences	5000 ms
Anomalous Sentences	6000 ms
Vowel Space	2000 ms
Rainbow Passage	1 minute
Spontaneous Speech	5 minutes

Table 3. Recording times for the different material sets.

A slightly different procedure was used for the disyllabic nonwords, the connected speech passage, and the spontaneous speech task. For the nonwords, all of the wS words were presented in random order before the Sw words, which were also presented in random order. On each trial, participants saw the carrier sentence with the nonword on the screen as well as the nonword in isolation and were instructed to read aloud the entire sentence and then to repeat the nonword by itself. Participants were given an example in the instructions and were asked to read the example aloud for the experimenter before beginning each nonword block. For the passage of connected speech, the participants were shown the passage along with the instructions and were asked to read it silently to themselves to familiarize themselves with it. They were then given 1 minute to read the passage aloud. The experimenter pressed the "s" key to stop the recording and continue on to the next experimental block when the participant had finished reading. For the spontaneous speech sample, the participant was asked to engage in a conversation with the experimenter. The experimenter asked the participant questions about her hometown and the participant was encouraged to speak freely for five minutes.

Looking Forward

Collection of the corpus has been completed for the 30 female talkers and our initial impression based on listening to the recordings is that we have captured some of the dialect variation we were looking for, despite the somewhat unnatural nature of the read speech materials. We expect to begin using these speech samples in the coming months in a number of novel perceptual and acoustic studies of the variation and variability of the English spoken in the state of Indiana.

In collecting this large corpus of speech under controlled laboratory conditions, we have gained valuable knowledge and insights into the kinds of read materials that are more or less likely to elicit variability in talkers from different regions. We are now working to streamline the recording process and to reduce the materials in order to make the project more feasible for extension to other parts of the country. In particular, we expect to include only half the number of sentences and multisyllabic words, a greatly reduced number of CVC words, and extended passage and spontaneous speech tasks in our larger study, the Nationwide Speech Project (NSP). The NSP will include recordings of talkers in Boston, New York City, Philadelphia, Atlanta, Chicago, Minneapolis, Dallas, Boulder, Los Angeles, and Seattle.

References

Baker, K., Ramig, L., Johnson, A., & Freed, C. (1997). Preliminary voice and speech analysis following fetal dopamine transplants in 5 individuals with Parkinson disease. *Journal of Speech, Language, and Hearing Research*, *40*, 615-626.

Bergquist, J.M. (1981). Tracing the origins of a midwestern culture: The case of central Indiana. *Indiana Magazine of History,* 1-32.

Byrd, D. (1992). Sex, dialects, and reduction. *ICSLP 92 Proceedings*, *92*, 827-830.

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, *15,* 39-54.

Carmony, M. (1965). Aspects of regional speech in Indiana. In L.M. Davis (Ed.), *Studies in Linguistics in Honor of Raven I. McDavid, Jr.* (pp. 9-24). Tuscaloosa, AL: University of Alabama Press.

Carter, A.K. & Clopper, C.G. (2000). Prosodic and morphological effects on word reduction in adults: A first report. In *Research on Spoken Language Processing Progress Report No. 24* (pp. 67-86). Bloomington, IN: Speech Research Laboratory, Indiana University.

Carver, C.M. (1987). *American Regional Dialects: A Word Geography*. Ann Arbor, MI: University of Michigan Press.

Cassidy, F. (1985). *Dictionary of American Regional English*. Cambridge, MA: Belknap Press of Harvard University Press.

Clopper, C.G. (2000). Some acoustic cues for categorizing American English regional dialects: An initial report on dialect variation in production and perception. In *Research on Spoken Language Processing Progress Report No. 24* (pp. 43-66). Bloomington, IN: Speech Research Laboratory, Indiana University.

Clopper, C.G. & Pisoni, D.B. (2002). Some new experiments on perceptual categorization of dialect variation in American English: Acoustic analysis and linguistic experience. Invited lecture at the 9th Biennial Rice Symposium on Linguistics, Houston, TX. March 14-16.

Cutler, A. & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133-142.

Darley, F.L., Aronson, A.E., & Brown, J.R. (1975). *Motor Speech Disorders*. Philadelphia: Saunders.

Davis, L.M. & Houck, C.L. (1992). Is there a Midland dialect? – Again. *American Speech, 67*, 61-70.

DuBois, J.W., Chafe, W.L., Meyer, C., & Thompson, S.A. (Eds). (2000). The Santa Barbara Corpus of Spoken American English. Santa Barbara, CA: University of California, Santa Barbara Center for the Study of Discourse.

Fairbanks, G. (1940). *Voice and Articulation Drillbook*. New York: Harper.

Fisher, W.M., Doddington, G.R., & Goudie-Marshall, K.M. (1986). The DARPA speech recognition research database: Specifications and status. *Proceedings of the DARPA Speech Recognition Workshop*, 93-99.

Gelfer, M. & Schofield, K. (2000). Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *Journal of Voice, 14*, 22-33.

Gerstman, L.H. (1968). Classification of self-normalized vowels. *IEEE Transactions of Audio and Electroacoustics AU-16*, 78-80.

Gibbens, V.E. (1962). Progress report on a word geography of Indiana. *Midwest Folklore, 11,* 151-154.

Hammond, M. (1999). *The Phonology of English*. Oxford: Oxford University Press.

Hayes, B. (1995). *Metrical Stress Theory*. Chicago: Chicago University Press.

Hillenbrand, J. & Houde, R. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39*, 311-321.

Kalikow, D.N., Stevens, K.N., & Elliot, L.L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America, 61*, 1337-1351.

Karl, J.R. & Pisoni, D.B. (1994). Effects of stimulus variability on recall of spoken sentences. In *Research on Spoken Language Processing Progress Report No. 19* (pp. 145-193). Bloomington, IN: Speech Research Laboratory, Indiana University.

Keating, P., Blankenship, B., Byrd, D., Flemming, E., & Todaka, Y. (1992). Phonetic analyses of the TIMIT corpus of American English. *ICSLP 92 Proceedings*, *92*, 823-826.

Keating, P.A., Byrd, D., Flemming, E., & Todaka, Y. (1994). Phonetic analyses of word and segment variation using the TIMIT corpus of American English. *Speech Communication*, *14*, 131-142.

Kelly, M. (1988). Phonological biases in grammatical category shifts. *Journal of Memory and Language, 27*, 343-358.

Kelly, M. & Bock, J. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 389-403.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change, 2,* 205-254.

Labov, W., Ash, S., & Boberg, C. (in press). *Atlas of North American English.* Mouton deGruyter.

Labov, W., Yeager, M., & Steiner, R. (1972). *A Quantitative Study of Sound Change in Progress.* Philadelphia: US Regional Survey.

Linguistic Data Consortium. (1996). CallFriend Telephone Speech Corpus for American English. (LDC96S46 & LDC96S47).

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1-36.

McHenry, M. (1999). Aerodynamic, acoustic and perceptual measures of nasality following traumatic brain injury. *Brain-Injury, 13*, 281-290.

Milroy, J. & Milroy, L. (1993). Mechanisms of change in urban dialects: The role of class, social network and gender. *International Journal of Applied Linguistics, 3,* 57-77.

Nusbaum, H.C., Pisoni, D.B., & Davis, C.K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception Progress Report No. 10* (pp. 357-376). Bloomington, IN: Speech Research Laboratory, Indiana University.

Peterson, G. & Barney, H. (1952). Control methods used in the study of vowels. *Journal of the Acoustical Society of America, 24*, 175-184.

Sapienza, C., Walton, S., & Murry, T. (1999). Acoustic variations in adductor spasmodic dysphonia as a function of speech task. *Journal of Speech, Language, and Hearing Research, 42*, 127-140.

Thomas, E.R. (2001). *An Acoustic Analysis of Vowel Variation in New World English*. Durham, NC: Duke University Press.

Torretta, G.M. (1995). The "easy-hard" word multi-talker speech database: An initial report. In *Research on Spoken Language Processing Progress Report No. 20* (pp. 321-334). Bloomington, IN: Speech Research Laboratory, Indiana University.

Zue, V., Seneff, S., & Glass, J. (1990). Speech database development at MIT: TIMIT and beyond. *Speech Communication*, *9*, 351-356.

Appendix 1

Repeated CVC Words

bean	bite	boat	boil
bull	calm	can	cap
caught	coal	code	coin
con	cool	cot	cough
death	dig	dime	dock
doll	doubt	dull	fade
fail	feed	fell	fire
fool	foul	full	gap
good	guide	head	heal
hill	home	keep	lit
loud	love	loyal	luck
lull	lung	main	math
meal	mile	mill	mob
pal	pen	pin	poke
pool	pull	rice	rip
sail	sell	sour	south
tape	tool	towel	town
tube	voice	void	walk
wall	wet	wool	wrong

Appendix 2

Disyllabic Nonwords

Sw Words		wS Words	
defloing	pagrawl	bockis	linnid
degroff	pecluff	bozen	lofek
deploze	pefrowl	cheelut	moggit
deprinn	recleal	chueliff	noilup
detrath	reglann	doolis	shullis
detrull	rekress	fautchep	soggith
diprool	retroud	fiden	taylup
fegroff	sebrull	foutik	teelub
fegroish	seclide	guidet	thellig
medrall	seglutch	hanneb	tullib
mitween	seplail	kowtiff	villut
nebrenn	sigrine	koysig	vullek
nebruel	tebroze	kulleff	whelid
neprale	teprill	lannis	yailub

Appendix 3

The Rainbow Passage (Fairbanks, 1940)

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for a pot of gold at the end of the rainbow.

IV. Publications

IV. Publications

ARTICLES PUBLISHED:

Carter, A. (2002). A phonetic and phonological analysis of weak syllable omissions: Comparing children with normally developing language and SLI. In E. Fava & A. Mioni (Eds.), *Issues in clinical Linguistics*. Padova, Italy: Unipress.

Carter, A. (2001). Featural morphology: Evidence from Muna irrealis affixation. *Coyote Papers Working Papers in Linguistics 10*, University of Arizona, 1-18.

Chin, S.B., Finnegan, K.R. & Chung, B.A. (2001). Relationships among types of speech intelligibility in pediatric users of cochlear implants. *Journal of Communication Disorders, 34,* 187-205.

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

Cleary, M. & Pisoni, D.B. (2002). Talker discrimination by prelingually-deaf children with cochlear implants: Preliminary results. *Annals of Otology, Rhinology, & Laryngology Supplement-Proceedings of the 8th Symposium on Cochlear Implants in Children, 111,* 113-118.

Cleary, M., Pisoni, D.B. & Geers, A.E. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear & Hearing, 22,* 395-411.

Clopper, C.G. (2001). The Nzema verbal phrase: An Optimality Theoretic account. *Indiana University Linguistics Club Working Papers Online 1(4)*.

Dinnsen, D.A., McGarrity, L.W. & O'Connor, K.M. (2001). Review of Rene Kager: Optimality Theory. *Studies in Second Language Acquisition, 23,* 422-423.

Dinnsen, D.A., McGarrity, L.W., O'Connor, K.M. & Swanson, K.A. (2000). On the role of sympathy in acquisition. *Language Acquisition*, *8,* 321-361.

Goh, W.D., Pisoni, D.B., Kirk, K.I., & Remez, R.E. (2001). Audio-visual perception of sinewave speech in an adult cochlear implant user: A case study. *Ear & Hearing, 22,* 412-419.

Harnsberger, J.D. (2001). On the relationship between identification and discrimination of non-native nasal contrasts: Evaluating the Perceptual Assimilation Model. *Journal of the Acoustical Society of America, 110,* 489-503.

Harnsberger, J.D. (2001). The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics, 29,* 303-327.

Harnsberger, J.D., Svirsky, M.A., Kaiser, A.R., Pisoni, D.B., Wright, R. & Meyer, T.A. (2001). Perceptual ‘vowel spaces’ of cochlear implant users: Implications for the study of auditory adaptation to spectral shift. *Journal of the Acoustical Society of America*, *109,* 2135-2145.

Herman, R. (2000). Phonetic markers of global discourse structures in English. *Journal of Phonetics,* 28, 466-493.

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

Miyamoto, R.T., Bichey, B.G., Wynne, M.K. & Kirk, K.I. (2002). Cochlear implantation with large vestibular aqueduct syndrome. *Laryngoscope, 112,* 1178-1182.

Vitevitch, M.S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 735-747.

BOOK CHAPTERS PUBLISHED:

Cleary, M. & Pisoni, D.B. (2001). Speech perception and spoken word recognition: Research and theory. In B. Goldstein (Ed.), *Handbook of Perception*. Cambridge: Blackwell. Pp. 499-534.

Frisch, S.A., Large, N.R., Zawaydeh, B. & Pisoni, D.B. (2001). Emergent phonotactic generalizations in English and Arabic. In J. Bybee & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. Pp. 159-180.

MANUSCRIPTS ACCEPTED FOR PUBLICATION (IN PRESS):

Bergeson, T.R. & Pisoni, D.B. (In press). Audiovisual speech perception in deaf adults and children following cochlear implantation. In G. Calvert, C. Spence & B.E. Stein (Eds.), *Handbook of Multisensory Integration*.

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

Carter, A. & Gerken, L. (In press). Similarities in weak syllable omissions between children with Specific Language Impairment and normally developing language: A preliminary report. *Journal of Communication Disorders*.

Cleary, M. & Pisoni, D.B. (In press). Speech Perception. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. London, UK: Macmillan.

Cleary, M., Pisoni, D.B. & Kirk, K.I. (In press). Working memory spans as predictors of word recognition and receptive vocabulary in children with cochlear implants. *Volta Review*.

Clopper, C.G. (In press). Frequency of Stress Patterns in English: A Computational Analysis. *Indiana University Linguistics Club Working Papers Online 2*.

Frisch, S., Broe, M. & Pierrehumbert, J. (In press). Similarity and phonotactics in Arabic. *Natural Language and Linguistic Theory.*

Goh, W.D. & Pisoni, D.B. (In press). Effects of lexical neighborhoods on immediate memory span for spoken words. *Quarterly Journal of Experimental Psychology.*

Herman, R. & McGory, J.T. (In press). The conceptual similarity of intonational tones and its effects on intertranscriber similarity. *Language and Speech*.

Herman, R. & Pisoni, D.B. (In press). Perception of "elliptical speech" following cochlear implantation: Use of broad phonetic categories in speech perception. *Volta Review*.

Kaiser, A.R., Kirk, K.I., Lachs, L. & Pisoni, D.B. (In press). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech-Language-Hearing Research.*

Lachs, L., McMichael, K. & Pisoni, D.B. (In press). Speech perception and implicit memory: Evidence for detailed episodic encoding of phonetic events. In J. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory.*

Lachs, L., Weiss, J.W. & Pisoni, D.B. (In press). Use of partial information by cochlear implant patients and normal-hearing listeners in identifying spoken words: Some preliminary analyses. *Volta Review*.

Meyer, T.A., Pisoni, D.B., Luce, P.A. & Bilger, R.C. (In press). An analysis of the psychometric and lexical neighborhood properties of spondaic words. *Journal of the American Academy of Audiology.*

Meyer, T.A., Pisoni, D.B., Luce, P.A. & Bilger, R.C. (In press). An analysis of the psychometric and lexical neighborhood properties of the spondaic words. *Journal of the American Academy of Audiology.*

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.*

Sheffert, S.M., Pisoni, D.B., Fellowes, J.M. & Remez, R.E. (In press). Learning to recognize talkers from natural, sinewave and reversed speech samples. *Journal of Experimental Psychology: Human Perception and Performance*.

Vitevitch, M.S., Pisoni, D.B, Kirk, K.I., Hay-McCutcheon, M. & Yount, S.L. (In press) Effects of phonotactic probabilities on the processing of spoken words and nonwords by postlingually deafened adults with cochlear implants. *Volta Review*.

Wong, D., Pisoni, D.B., Learn, J., Gandour, J.T., Miyamoto, R.T. & Hutchins, G.D. (In Press). Differential cortical activation to monaural speech and nonspeech stimuli: A PET imaging study. *Hearing Research*.

SUBMITTED:

Burkholder, R.A. & Pisoni, D.B. (Submitted). Speech timing and working memory in profoundly deaf children after cochlear implantation. *Journal of Experimental Child Psychology.*

Carter, A. & Clopper, C. (Submitted). Word reduction by adults: Effects of prosodic patterns in the input. *Language and Speech*.

Carter, A. & Gerken, L. (Submitted). Do children's omissions leave traces? *Journal of Child Language.*

Clopper, C.G. & Pisoni, D.B. (Submitted). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*.

Clopper, C.G. & Pisoni, D.B. (Submitted). Some effects of talker variability on perceptual learning of American English regional dialects. *Psychological Science*.

Clopper, C.G & Pisoni, D.B. (Submitted). Homebodies and Army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change.*

Cummings, K.E., Chin, S.B. & Pisoni, D.B. (Submitted). Peturbation of speech after consumption of alcohol. *Journal of the Acoustical Society of America*.

Dillon, C.M., Burkholder, R.A., Cleary, M. & Pisoni, D.B. (Submitted). Perceptual ratings of nonword repetition responses by deaf children after cochlear implantation: Correlations with measures of speech, language and working memory. *Journal of Speech, Language and Hearing Research*.

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

Ferguson, S.H. & Kewley-Port, D. (Submitted). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 112,* 259-271.

Harnsberger, J.D., Wright, R. & Pisoni, D.B. (Submitted). Effects of speaking style on the perceptual learning of novel voices: A first report*.* *Phonetica*.

Houston, Pisoni, D.B., Kirk, K.I., Ying, E.A. & Miyamoto, R.T. (Submitted). Speech perception skills of deaf children following cochlear implantation: A first report. *International Journal of Pediatric Otolaryngology.*

Karpicke, J. & Pisoni, D.B. (Submitted). Multimodal effects on immediate memory span. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Karpicke, J. & Pisoni, D.B. (Submitted). Using immediate memory span to measure implicit learning. *Psychonomic Bulletin & Review*.

Lachs, L. & Pisoni, D.B. (Submitted). Effects of multi-modal speech cues on recognition memory for spoken words. *Journal of Experimental Psychology: Human Perception and Performance.*

Lachs, L. & Pisoni, D.B. (Submitted). Visual recognition of spoken words without audition. *Perception & Psychophysics.*

Lachs, L. & Pisoni, D.B. (Submitted). Crossmodal source information and spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance.*

Lachs, L. & Pisoni, D.B. (Submitted). Crossmodal source identification in speech perception. *Perception & Psychophysics*.

Lachs, L. & Pisoni, D.B. (Submitted). Specification of crossmodal source information in isolated kinematic displays of speech. *Journal of the Acoustical Society of America.*

McMichael, K.H. & Pisoni, D.B. (Submitted). Effects of talker-specific encoding on recognition memory for spoken sentences. *Memory & Cognition*.

Pisoni, D.B., & Cleary, M. (Submitted). Measures of working memory span and verbal rehearsal speed in deaf children following cochlear implantation. *Ear & Hearing*.

Pisoni, D.B., Kirk, K.I., Svirsky, M.A. & Miyamoto, R.T. (Submitted). Looking at the “Stars”: A first report on the intercorrelations among measures of speech perception, intelligibility and language development in pediatric cochlear implant users. *Journal of Speech, Language and Hearing Research.*

Roediger, H.L., III, McDermott, K.B., Pisoni, D.B. & Gallo, D.A. (Submitted). Illusory recollection of voices. *Memory.*

Vitevitch, M.S. (Submitted). Change Deafness: The inability to detect changes in a talker's voice. *Journal of Experimental Psychology: Human Perception and Performance.*

Vitevitch, M.S. (Submitted). Influences of the number of phoneme positions that form neighbors on spoken word recognition. *Psychonomic Bulletin & Review*.

Vitevitch, M.S. & Sommers, M. (Submitted). The role of phonological neighbors in the tip-of-the-tongue state. *Memory and Cognition*.