INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9119626

Paired-comparison judgments for hearing aid selection in children

Eisenberg, Laurie Sue, Ph.D.

City University of New York, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

PAIRED-COMPARISON JUDGMENTS
FOR HEARING AID SELECTION IN CHILDREN

By

LAURIE SUE EISENBERG

A dissertation submitted to the Graduate Faculty in Speech and Hearing Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1991

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Chair of Examining Committee

October 17, 1990
Date

Irving Hochberg
Executive Officer

Harry Levitt
Arthur Boothroyd
Irving Hochberg
Arlene Neuman Sand
Donald W. Nielsen

Supervisory Committee

The City University of New York

Abstract

PAIRED-COMPARISON JUDGMENTS
FOR HEARING AID SELECTION IN CHILDREN

By

Laurie Sue Eisenberg

Adviser: Professor Harry Levitt

Three experiments were conducted to investigate the feasibility of using a paired-comparison technique to select hearing aids for children. Experiment 1 determined the age at which 5 out of 5 normal-hearing children were able to pass a paired-comparison test involving judgments of auditory clarity. Twenty-five normal-hearing children between the ages of 4.0 and 6.5 years were tested. Five age groups were formed with 5 subjects per group. The results showed that all 5 children between the ages of 6.0 and 6.4 years passed the auditory paired-comparison test.

Experiment 2 determined the age at which 5 out of 5 hearing-impaired children were able to pass a paired-comparison test involving judgments of auditory clarity. Ten children with mild to moderately-severe sensorineural hearing loss were tested. Their ages ranged between 5.5 and 7.4 years. Two age groups were formed with 5 subjects per group; 5.5 to 6.4 years, and 6.5 to 7.4 years. The results showed that all 5 hearing-impaired children between the ages of 6.5 and 7.4 years passed the auditory paired-comparison test.

Experiment 3 investigated whether a paired-comparison technique could be used effectively to select hearing aids for hearing-impaired children capable of performing an auditory paired-comparison task. Eight hearing-impaired children between the ages of 5.7 and 7.8 years judged the clarity of seven hearing aid, frequency-gain shapes in two paired-comparison tournaments. Phoneme identification scores were also obtained for each shape.

The results showed that correlations between the two tournaments were moderate to strong for 6 children and weak for 2 children. Of the 6 children who demonstrated good reliability, correlations between paired-comparison judgments and phoneme identification scores were moderately strong for 3 children and weak to moderate for 3 children. The hearing aid selected by the paired-comparison technique did not degrade speech intelligibility when compared with a standard prescription. Results from this study indicate that a paired-comparison technique can be used reliably to select hearing aids for some hearing-impaired children who are 6.5 years of age and occasionally younger.

ACKNOWLEDGMENTS

This study was funded in part by the Confidence Foundation and the Guggenheim Foundation. I wish to thank the House Ear Institute and the following staff members who contributed to this project in a substantial way: Lars Nielsen, Ed Wu, Ross Adams, Carolyn Brown, Janet Stoeckert, Michael Nilsson, Karen Berliner, Lisa Tonokawa, Randy Green, and Butch Welch.

To my committee -- Harry Levitt, Arthur Boothroyd, Irving Hochberg, Arlene Neuman Sand, and Donald Nielsen -- thank you for your expert guidance, patience, and kindness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Sensorineural hearing loss in children warrants immediate and continued attention because of the important contribution of auditory information to the development of speech and language. A child's native language is acquired early in life primarily through the auditory channel. Early onset of hearing loss impedes this development, creating an urgency for remediation. Delayed or ineffective management can impact on the cognitive, emotional, and social functioning of the hearing-impaired child.

At the present time, sensorineural hearing loss is medically untreatable. The most viable alternative is prosthetic intervention, most notably amplification by electroacoustic hearing aids. As Ross and Tomassetti (1980) have stated, "for most hearing-impaired children, the early and appropriate selection and use of amplification is the single most important habilitative tool available to us" (p. 214). Unfortunately, appropriate selection of hearing aids is not a trivial matter.

Hearing aid selection techniques presently used with children are modified from standard methods developed for adults. Though seemingly adequate, such techniques are restricted by age-related constraints in testing. Recently, more flexible testing techniques have been explored and found advantageous in selecting hearing aids for adults. The potential for improved hearing aid selection justifies investigation of these new techniques for children.

This study investigates the feasibility of using one such technique, paired-comparison judgments. Hearing aid conditions are paired systematically and evaluated until one condition emerges as the preferred choice. This chapter provides relevant background information to support investigation of the paired-comparison technique with children. The discussion concludes with a summary detailing the aims and design of the study.

Review of the Literature

The traditional methods used to select hearing aids for adults and children are explored in this review. A more advantageous technique is described in which adults select hearing aids by paired-comparison judgments. The feasibility of using a paired-comparison technique with children is discussed in relation to the linguistic and cognitive abilities required to make comparative judgments.

Hearing Aid Selection

Hearing aid selection techniques aim to amplify speech in a way that will maximize intelligibility without exceeding the threshold of discomfort. Different approaches have been promoted to achieve this goal. Two, the comparative and prescriptive approaches, have gained widespread clinical acceptance and are described below.

The comparative approach evaluates the patient's performance with different hearing aids (Carhart, 1946). Test measures include the speech reception threshold, speech discrimination score, tolerance limits for loud sounds, and speech discrimination in noise. The hearing aid providing the best speech discrimination and the widest dynamic range is the one typically selected. Because patients get experience with actual hearing aids, subjective factors such as comfort, pleasantness, and quality of sound also play a role in the decision-making process (Green, Day, & Bamford, 1989; Levitt, 1978a).

Although it is appealing to compare performance among hearing aids, there are four disadvantages to this approach. First, testing with more than one hearing aid is very time consuming and fatiguing for the patient. Secondly, there is no guarantee that the optimal hearing aid will be selected because there are no specific guidelines for choosing aids to compare (Levitt, 1978a;

Pascoe, 1985). That is, if the set of aids chosen for evaluation does not contain an appropriate hearing aid for the patient, then the result will be far from optimal despite extensive testing.

Thirdly, the clinician does not have a measurable goal that can be verified by precise physical measurements. Another negative result is that, despite the large number of hearing aids fitted this way, there is no attempt (because of the underlying philosophy) to determine the electroacoustic characteristics that identified the best hearing aid.

Lastly, and most important, standard speech discrimination word lists are insensitive and unreliable as measures used to differentiate hearing aids. Although Carhart (1946) considered an 8% difference in performance to be meaningful, Thornton and Raffin (1978) demonstrated that such differences could occur by chance with a probability of .32 for two 50-item word lists even though the true test scores for the two word lists both equal 50%.

A study conducted by Shore, Bilger, and Hirsh (1960) looked at the reliability of standard speech tests in selecting hearing aids. Spondees and 25-item monosyllabic word lists (presented in quiet and noise) were used to evaluate four different commercial hearing aids. Fifteen hearing-impaired adults were tested with each aid four times. Results revealed that test scores varied across sessions for each subject. Shore et al. (1960) concluded that existing methods of speech audiometry were not sufficiently sensitive to measure performance differences among hearing aids.

In another study, Walden, Schwartz, Williams, Holum-Hardegen, and Crowley (1983) assessed performance among different hearing aids using 100-item monosyllabic word lists in the presence of multitalker babble. They found that word discrimination scores differed significantly only when the hearing aids being evaluated exhibited substantial electroacoustic differences.

A second, possibly more serious finding emerged from the Walden et al. (1983) study. Following a trial period with three electroacoustically-different hearing aids, the hearing aid that scored highest on speech discrimination tests was shown to be the most preferred in only 50% of the cases.

The alternative to the comparative approach is hearing aid fitting by prescription. Standard audiometric measures provide the information needed to select an aid's electroacoustic characteristics of gain, frequency response, and maximum power output. In contrast to the comparative approach, the prescriptive approach provides the clinician with a goal that can be verified by real-ear measures.

A central objective of the prescriptive approach is to adjust gain as a function of frequency so as to make the elements of speech more audible and intelligible to the hearing-impaired listener. In fitting a hearing aid, the clinician strives to amplify speech to a most comfortable loudness level without exceeding the listener's loudness discomfort level.

Prescriptive techniques have been advocated by a number of investigators, each proposing specific rules or formulas for determining real-ear gain, frequency response, and maximum power output. To determine the frequency-gain characteristic, some prescriptions incorporate puretone thresholds into the formula (Berger, 1976; Byrne & Dillon, 1986; Byrne & Tonisson, 1976; Libby, 1985; Lybarger, 1978; McCandless & Lyregaard, 1983). Other prescriptions utilize suprathreshold information such as most comfortable loudness level (Pascoe, 1978; Shapiro, 1976; Skinner, Pascoe, Miller, & Popelka, 1982), or upper level of comfort (Cox, 1983). The loudness discomfort level is required in some prescriptions to select the maximum power output (Berger, 1976; McCandless & Lyregaard, 1983; Shapiro, 1976).

The threshold-based prescriptions are easier to implement and are widely used. Suprathreshold-based prescriptions promise a more individualized hearing aid fitting but require additional measurements which are time consuming and may introduce sources of variability. Furthermore, suprathreshold tests are difficult to accomplish with hard-to-test populations.

Prescriptions also differ in the way they are calculated. Some prescriptions apply simple mathematical corrections to threshold information (Berger, 1978; Lybarger, 1978; McCandless & Lyregaard, 1983). Other prescriptions, more theoretically-based, incorporate corrections for the long-term average speech spectrum (Byrne & Dillon, 1986; Byrne & Tonisson, 1976; Cox, 1983; Skinner et al., 1982), the slope of the audiogram (Byrne & Dillon, 1986), and the equal-loudness comfort level contour (Byrne & Tonisson, 1976; Watson & Knudsen, 1940).

A number of studies have compared the various prescription methods (Collins & Levitt, 1980; Leijon, Eriksson-Mangold, & Bech-Karlsen, 1984; Lippmann, Braida, & Durlach, 1981; Pascoe, 1975; Skinner, 1980; Sullivan, Levitt, Hwang, & Hennessey, 1988). Those studies confirmed that different prescriptions yielded different frequency-gain response curves, but no one formula produced superior results across studies.

In one study, Sullivan et al. (1988) compared four standard prescription methods at three different output levels. Fourteen hearing-impaired adults were tested on word recognition and judgments of relative speech intelligibility and quality. Results revealed that performance by prescription varied as a function of output level although mean differences were small. In addition, mean differences in frequency-gain curves became negligible when output level was adjusted for listener comfort. On the other hand, performance by prescription varied for individual subjects, suggesting that no one prescription can account for the individual differences observed among the hearing impaired.

In another study, Skinner (1980) compared subject performance on frequency responses that increased gain systematically in the high frequencies. Six subjects with sloping, high-frequency hearing losses were tested at different input levels on word lists weighted with high-frequency consonants. For the high gain conditions, the best performance was achieved at low input levels, whereas for the low gain conditions, the best performance was achieved at higher input levels. Based on these findings, Skinner (1980) surmised that a single, fixed-frequency response would not adequately compensate for the varying input levels encountered in the environment.

To summarize, hearing aid selection techniques include the comparative and prescriptive approaches. Although both approaches are widely used, the prescriptive approach offers the greater advantage by providing a verifiable electroacoustic target. However, it is also evident that no one prescription is optimal for all hearing aid users in all situations.

Although much effort has been devoted to hearing aid selection, present-day clinical techniques were developed primarily for adults. It is even more crucial that appropriate amplification be selected for children because of the important contribution of hearing to communication development. Techniques developed for use with children have been modified from adult procedures. The effectiveness of these techniques must be weighed in relation to the developmental constraints that restrict the testing of children.

Hearing Aid Selection for Children

Hearing aid selection techniques presently used with children have been adapted from comparative and prescriptive approaches. As noted above, both approaches rely heavily on standard audiologic measures. Such requirements impose necessary restrictions where children are concerned.

The audiometric test battery is subject to age-related constraints, and testing is limited to behavioral and objective threshold measures. Behavioral thresholds have been shown to vary over time (Fior, 1972; Elliott & Katz, 1980a; Maxon & Hochberg, 1982; Siegenthaler, 1969). For example, Maxon and Hochberg (1982) compared puretone thresholds for normal-hearing children aged 4, 6, 8, 10, and 12 years. Thresholds were seen to improve systematically with increasing age. The 12-year-old children exhibited thresholds slightly poorer than those reported for adults. It is not clear, however, whether the lower measured thresholds were due to maturation of the auditory system or to children's cognitive skills.

With use of modified test procedures, consistent behavioral thresholds have been obtained for very young children. Infants as young as 5 months have been shown to respond consistently to auditory stimuli using a head-turn response and visual reinforcement (Moore & Wilson, 1978). Using this procedure, Nozza and Wilson (1984) found that infants at 6 and 12 months of age exhibited poorer puretone thresholds than did adults tested at 1000 and 4000 Hz. The gap was quite narrow (5 to 7 dB) at 4000 Hz. These findings demonstrated that differences in auditory thresholds between infants and adults are smaller than was once believed.

Objective measures, such as the acoustic reflex or auditory brainstem response, provide an alternative to behavioral tests. However, these tests present difficulties in obtaining frequency-specific auditory information (Hall & Ruth, 1985). More recently, otoacoustic emissions have been shown to be a more effective means of obtaining this type of information in normal-hearing infants (Bonfils, Avan, Francois, Marie, Trotoux, & Narcy, 1990). However, the technique is less effective for ears exhibiting audiometric thresholds poorer than 30 dB HL.

With respect to speech recognition measures, children's ability to understand speech develops over time. This development has been attributed to a general maturation of perceptual, cognitive, and language- and information-processing abilities (Jerger, 1984). Age-related differences in performance were observed by Jerger (1984), who compared performance-intensity (PI) functions of six pediatric word tests for normal-hearing children 3 to 6 years of age. The PI functions were found to be 2 to 4 dB steeper for the older children.

Boothroyd (1968a, 1970) observed age-related differences in maximum word discrimination score for children between the ages of 6 and 9.5 years. He suggested that the lower scores obtained by the younger children could be attributed to an underdeveloped set of recognition categories for words and/or phonemes, articulation deficits, and an inability to use contextual information in making decisions.

Hearing aid selection procedures in use today vary and are dependent on the child's age at onset of hearing loss, the type and degree of the loss, and the age at time of diagnosis. Matkin (1986) and Northern and Downs (1984) advocate a comparative approach composed of tests that measure auditory sensitivity for puretones and speech, discrimination in quiet and noise, and tolerance for loud sounds.

Matkin (1986) described hearing aid selection as a "process by which the audiologist chooses one or two hearing aids with maximum electroacoustic flexibility from the available clinic stock to be tried during assessment of aided performance" (p. 178). The hearing aids selected are often based on clinician insight and past experience with older children or adults. Unfortunately, this method will present a problem with the new digital hearing aids because they require so many variables to be adjusted.

Matkin (1986) also provided guidelines for selecting the maximum power output. He recommended that no hearing aid output exceed 125 dB SPL for children with profound losses; 120 dB SPL for moderate losses; and 110 dB SPL for mild losses. With preschool-aged children, it was suggested that an additional 5 dB be subtracted from the above levels.

Selection procedures recommended by Northern and Downs (1984) were categorized according to the age and language abilities of the child. For infants and young preverbal children, informal observation of sound awareness was recommended in estimating gain and maximum power output. Three or four hearing aids were selected for evaluation. Speech discrimination measures were incorporated with children 3 years of age and older, and standard adult procedures were considered appropriate for children between the ages of 10 and 16 years. The speech discrimination score was given the greatest weight in the selection process.

Because standard speech discrimination tests were considered to be inappropriate for children, new tests were developed specifically for this population. Matkin (1986) recommended that children's tests be administered according to the child's receptive vocabulary age (RVA). He advised that nonverbal children be tested on the Ling 5-Sound Test (Ling & Ling, 1978) and phoneme imitation. Tests appropriate for children with an RVA up to 4 years included the Sound Effects Recognition Test (Finitzo-Hieber, Gerling, Matkin, & Cherow-Skalka, 1980) and the Auditory Numbers Test (Erber, 1980). For children with an RVA of 4 to 6 years, the Word Intelligibility by Picture Identification (WIPI) Test (Ross & Lerman, 1970) and the Northwestern University Children's Perception of Speech (NU-CHIPS) Test (Elliott & Katz, 1980b, 1980c) were recommended. For an RVA of 6 to 12 years, the phonetically balanced PBK-50 word lists (Haskins, 1949) were advocated. Standard adult word lists were considered to be appropriate for children 12 years of age and older.

Speech discrimination tests developed for young children have been criticized as insensitive and unreliable. Resnick (1984) observed that most tests developed for young children are closed set and use pictures to represent words. She pointed out that the number of words available for this type of test is limited by children's vocabulary and by words that can be represented in pictures. She also noted that some tests reduced the number of items or response alternatives to compensate for limited attention spans. Although such modifications were deemed necessary to accommodate children's abilities, Resnick (1984) concluded that such alterations must lessen the usefulness of speech tests in selecting hearing aids.

One technique that compares performance among hearing aids but does not rely on speech tests is the aided trials approach (Schwartz & Larson, 1977b; Studebaker, 1982). This approach compares aided soundfield thresholds for different hearing aids and/or settings. The hearing aid (or setting) that provides the greatest improvement in threshold and the smoothest response across frequencies is the one typically selected.

Although the aided trials approach is widely used, several weaknesses have been noted. First, because the aided soundfield audiogram is limited to threshold information, there is no indication of performance at suprathreshold levels (Schwartz & Larson, 1977b). Secondly, the aided trials approach does not specify an optimal target, and there is the risk of providing too much gain in an attempt to bring thresholds close to normal (Studebaker, 1982).

Thirdly, soundfield thresholds have been shown to be variable. Hawkins, Montgomery, Prosek, and Walden (1987) calculated critical differences for aided soundfield thresholds at different frequencies and probability levels for a group of hearing-impaired adults. They showed that two aided thresholds at 1000 Hz would have to differ by 15 dB or more to ensure that the difference was significant at the .05 probability level. They also commented that aided thresholds can be affected by internal noise of the hearing aid and/or ambient noise in the test environment. It should be noted that smaller critical differences (e.g., 10 dB at 1000 Hz) were reported by Stuart, Durieux-Smith, and Stenstrom (1990) for a group of hearing-impaired children between the ages of 5 and 14 years.

Hearing aid prescriptions have also been adapted for use with children. Frequency-gain settings are first estimated from auditory threshold data, becoming more finely tuned as the child gains experience with sound. As noted with adults, the electroacoustic characteristics of the aid can be verified with real-ear measures or aided/unaided soundfield measures of functional gain (Beauchaine, Barlow, & Stelmachowicz, 1990).

One early prescription method originated from a suprathreshold-based procedure. Developed by Gengel, Pascoe, and Shore (1971), the hearing aid selected was the one that best amplified the average speech spectrum to audible levels. According to this procedure, the child first adjusted the volume control of the hearing aid to a preferred listening level while recorded speech was presented at an input level of 70 dB SPL. At this setting, aided thresholds were obtained using narrow bands of noise. The aided thresholds were subtracted from noise-band levels computed from average speech levels, and the difference represented the sensation level at which conversational speech was perceived. The aim of this approach was to select a hearing aid that amplified the speech spectrum 10 to 20 dB above the aided threshold.

According to Gengel et al. (1971), if some of the noise bands were not amplified above threshold, then a different aid was evaluated or the child would be trained to wear the present aid at a higher volume. Conversely, if some of the noise bands were seen to drive the aid beyond the maximum linear output (as measured by a bracketing technique with the aid turned full-on), then they recommended that the aid not be used at a full-on setting.

The Gengel et al. (1971) procedure was later modified by Erber (1973) to provide a more direct estimate of the sensation level of speech. According to this procedure, speech-spectrum noise bands (or filtered bands of real speech) were routed through a loudspeaker to the hearing aid while it was worn by the child at normal use setting and at several tone and volume settings. The hearing aid receiver was coupled to a sound level meter, and the output was compared with unaided noise-band thresholds measured through an insert receiver attached to the child's earmold. The difference between the aided and unaided levels represented the sensation level of speech as measured in the ear canal.

The hearing aid (and setting) that provided the best acoustic output across frequencies (10-25 dB above threshold and 5 dB below the maximum linear acoustic output) was the one selected. Although Erber (1973) suggested rejecting any hearing aid in which the maximum linear output exceeded the child's loudness discomfort level, there was no explanation of how the discomfort level was measured.

Schwartz and Larson (1977a) compared the Gengel et al. (1971) and Erber (1973) procedures with the aided trials approach in 10 hearing-impaired children between the ages of 5 and 10 years. In this study, each child was evaluated on the three procedures with his or her own hearing aid and two stock aids. Results indicated that the three procedures produced the same recommendations for children exhibiting mild to moderate hearing losses. For children with severe to profound losses, however, the aided trials approach was seen to overestimate the potential benefit of the hearing aid. That is, the hearing aid providing the best aided thresholds was not always adequate in amplifying the speech spectrum to optimal levels when re-evaluated by the speech-spectrum approach.

Most recently, Seewald and Ross with colleagues devised a means of prescribing the "desired sensation level" for speech from children's pure-tone thresholds (Ross & Seewald, 1988; Seewald & Ross, 1988; Seewald, Ross, & Spiro, 1985; Seewald, Ross, & Stelmachowicz, 1987). This procedure estimates a frequency-gain target that amplifies speech to audible levels, which can then be verified by real-ear measures. The procedure also accounts for the physical differences between a child's ear and adult's ear. In addition, the maximum power output (SSPL90) is predicted by plotting the difference between the amplified speech level and the recommended SSPL90 as a function of the projected level of amplified speech.

Of late, researchers have set out to expand the pediatric test battery by attempting to better quantify the dynamic range. Kawell, Kopun, and Stelmachowicz (1988) adapted the Hawkins, Walden, Montgomery, and Prosek (1987) loudness discomfort level (LDL) procedure for use with hearing-impaired children. In the Kawell et al. (1988) study, pictorial descriptors were used to represent five loudness categories (ranging from "too soft" to "hurts"). Children were required to point to the corresponding picture while warble-tone stimuli were presented in 2-dB increments from a loudspeaker. Tests were conducted with the child wearing a high-output hearing aid. The measures were repeated three times within a one- to three-week period. Results indicated that hearing-impaired children between the ages of 7 and 14 years were able to perform the task reliably.

Using a different approach, Macpherson, Elfenbein, Schum, and Bentler (1989) trained the concept of "too much" in several tasks to obtain LDLs in normal-hearing children between the ages of 3 and 5 years. In one task, the tester filled a cup with water and the child was instructed to push a button to indicate when the cup was filled. This concept was gradually transferred to warble tones that increased in intensity. Results between test and retest revealed that 5-year-old children could produce reasonably reliable LDLs. These two studies demonstrate that children as young as 5 years of age are capable of being trained to perform certain auditory tasks that were once believed to be too difficult.

To conclude, hearing aid selection techniques used with children have been modified from standard procedures because of constraints in testing. Recent studies have shown that children as young as 5 years of age are capable of performing tasks that measure LDLs. Other, more flexible techniques would also be advantageous for children. One technique, paired-comparison judgments, has been explored with adults and offers a number of advantages. The strengths and weaknesses of the paired-comparison technique are described below.

Paired-Comparison Judgments

The paired-comparison technique uses the listener's relative judgments of preference, quality, intelligibility, or other relevant attributes to select a hearing aid. Hearing aid conditions are paired systematically and evaluated by these relative judgments. Some form of tournament or adaptive procedure is used until the condition yielding the best results (according to the judgmental criteria) is identified. With this technique, any hearing aid characteristic can be evaluated on any perceptual dimension of sound quality using any type of auditory stimulus. In most studies, hearing aid conditions have been judged on intelligibility and quality of sound. Continuous discourse (with or without background noise) has typically been the stimulus of choice.

In the past, technical limitations precluded widespread clinical application of the paired-comparison technique. Early investigators tape recorded pairs of hearing aid processed speech in predetermined order for presentation to the listener (Punch, 1978; Punch & Beck, 1980; Punch & Howard, 1978; Punch, Montgomery, Schwartz, Walden, Prosek, & Howard, 1980; Punch & Parker, 1981; Studebaker, Bisset, Van Ort, & Hoffnung, 1982; Witter & Goldstein, 1971; Zerlin, 1962). Other researchers employed programmable master hearing aids to manipulate adaptively the characteristic being evaluated (Levitt, Sullivan, Neuman, & Rubin-Spitz, 1987; Neuman, Levitt, Mills, & Schwander, 1987; Sullivan et al., 1988). Now, digital hearing aid systems are commercially available and offer a number of programming capabilities for clinical use. Two of these systems provide software for paired-comparison testing (Hecox & Punch, 1988; Pluvinage & Benson, 1988).

Three types of paired-comparison procedures have been explored; the round-robin tournament, the single- and double-elimination tournament, and the simplex method. Each is briefly described.

The round-robin tournament pairs each hearing aid with every other hearing aid to form n(n-1)/2 pairs for a set of n hearing aids. Hearing aids are ranked according to the number of times each is selected. Early studies using tape recordings of processed speech often employed a round-robin procedure. More recently, an adaptive or "iterative" round-robin procedure was implemented by Levitt, Sullivan, Neuman, & Rubin-Spitz (1987) and Neuman et al. (1987) using programmable master hearing aids.

According to this procedure, an initial estimate of the optimum hearing aid is compared with conditions that vary systematically from the initial one. The winning condition becomes the initial estimate in a second round, and the tournament continues until one condition is ranked the winner in successive rounds.

Single- and double-elimination tournaments have also been used adaptively to select hearing aids by paired-comparison judgments (Neuman et al., 1987; Studebaker et al., 1982; Sullivan et al., 1988). Elimination tournaments advance winners of paired matches until one succeeds to first place. In the single-elimination tournament, only the winners continue on in the competition. The double-elimination tournament additionally matches the losers and eventually pairs the "winner of the winners" with the "winner of the losers" in the final match.

Because more trials per condition are required in the double-elimination tournament, it is the more reliable procedure (Montgomery, Schwartz, & Punch, 1982). However, elimination tournaments in general are subject to variability. Montgomery et al. (1982) demonstrated that the final outcome of an elimination tournament can be influenced by the arrangement of pairs at the outset.

The simplex method is the most complex of the adaptive paired-comparison procedures (Levitt, 1978a, 1978b; Levitt, Sullivan, Neuman, & Rubin-Spitz, 1987; Neuman et al., 1987). Derived from multivariate statistics (Box, 1957), the listener's responses determine the direction of adjustment along the one or more dimensions being manipulated. The initial estimate of the optimum hearing aid is first compared with conditions that deviate slightly from the initial one. The conditions continue to vary in the direction of preference until the listener's responses converge on the selected condition.

In a comparative study, Neuman et al. (1987) evaluated the iterative round-robin tournament, double-elimination tournament, and simplex method. Eight hearing-impaired adults judged the intelligibility of continuous discourse as it varied in frequency-gain shape. The three procedures produced similar outcomes but differed in the amount of time required to complete each test. On the average, the simplex method required 8 minutes to complete; the double-elimination tournament required 36.3 minutes to test 16 conditions; and the round-robin procedure required 83.8 minutes to complete two tournaments of nine hearing aids each.

The paired-comparison technique can be viewed as a particularly efficient version of the comparative approach. In a study conducted by Studebaker et al. (1982), a judgment between two aided conditions required approximately 20 to 40 seconds to complete, and an elimination tournament required 20 to 25 minutes to compare six or eight hearing aids; half of that time involved switching tapes. In contrast, the amount of time needed to administer speech discrimination tests for the same number of aids ranged from 35 to 45 minutes.

Studebaker et al. (1982) also presented data to suggest that the paired-comparison technique is the more reliable of the two methods. In three elimination tournaments, 15 normal-hearing and 10 hearing-impaired subjects judged intelligibility of continuous discourse in noise processed by 23 hearing aids. Speech discrimination tests were also administered for each aided condition. All tests were repeated. For normal-hearing subjects (tested at 0 dB signal/noise), paired-comparison judgments were more reliable than were discrimination scores, particularly when the range of discrimination scores was narrow. For hearing-impaired subjects (tested at +7 dB signal/noise), test-retest reliability was better for paired-comparison judgments than for discrimination scores. However, at 0 dB signal/noise, reliability was comparable for the two measures.

The top two winning hearing aids in each of the three elimination tournaments went on to compete in a round-robin tournament. Test-retest rankings produced mean correlation coefficients of .69 for the normal-hearing subjects and .64 for the hearing-impaired subjects. In contrast, test-retest rankings for corresponding speech discrimination scores produced mean coefficients of .31 for the normal hearing and .37 for the hearing impaired.

The reliability of paired-comparison judgments has varied across studies depending on the population tested and/or stimulus presented. For example, in a study conducted by Punch and Howard (1978), 90 normal-hearing subjects judged intelligibility of speech in quiet and multitalker babble (-7 dB signal/noise) processed by five hearing aids. Correlations between test and retest revealed coefficients of .94 for the quiet condition and .65 for the babble condition.

In another study conducted by Punch (1978), 10 normal-hearing and 10 hearing-impaired adults judged quality of male-spoken continuous discourse, female-spoken continuous discourse, and music. Each stimulus condition was processed by five hearing aids. The number of preferences were pooled for each condition per group. With male speech, test-retest correlations were essentially the same for each group; r = .86 for normal hearing, and $r = .85$ for hearing impaired. However, with female speech, test-retest reliability decreased and the discrepancy between groups increased; $r = .69$ for normal hearing, and $r = .54$ for hearing impaired. With music, test-retest reliability decreased further and the discrepancy between groups widened; $r = .56$ for normal hearing, and $r = .35$ for hearing impaired.

When comparing performance between groups, Punch (1978) found that the hearing-impaired subjects were more apt to elicit ties in response patterns, and he surmised that the hearing impaired were less sensitive to small differences among stimulus conditions. Punch and Parker (1981) noted that an inability to perceive differences among conditions contributed to increased test-retest variability. Levitt, Sullivan, Neuman, and Rubin-Spitz (1987) observed the occurrence of "circularities" (e.g., X>Y>Z>X) in listener judgments, and they speculated that this type of pattern could result from random guessing or variations over time in the judgmental criteria used by the listener.

Paired-comparison judgments may also be influenced by order effects and/or memory. In a study that assessed hearing aid quality, Green et al. (1989) found that subjects generally selected the second of two hearing aid settings following 1 minute of listening per setting.

This was true even when the settings being compared were identical. It also has been noted that the reliability of listeners' judgments is poor when a period of time elapses between comparisons (Hawkins, 1985).

Because the paired-comparison technique relies on subjective judgments, listeners may base judgments on some self-defined criteria rather than on criteria specified by test instructions. To test for this possibility in judgments of relative intelligibility, researchers have measured the relationship between paired-comparison judgments of intelligibility and scores obtained on speech recognition tests. In the study described earlier, Studebaker et al. (1982) demonstrated that the top two hearing aids in each elimination tournament also yielded the highest mean discrimination scores for the normal-hearing subjects. The average correlation coefficient between the round-robin rankings and discrimination score rankings was .77 (at 0 dB signal/noise). Moreover, confidence ratings assigned to judgments were correlated with the magnitude of difference between discrimination scores. Results revealed that discrimination score differences of 8% or greater corresponded highly with the confidence ratings assigned to those judgments.

Hearing-impaired subjects did not perform as well as normal-hearing subjects; tournament winners did not always yield the highest discrimination scores. Correlations between round-robin rankings and discrimination score rankings produced average coefficients of .30 (at 0 dB signal/noise) and .48 (at +7 dB signal/noise). Differences greater than 10% to 12% in discrimination score were associated with the confidence ratings assigned to the judgments.

In another study of this type, Levitt, Sullivan, Neuman, and Rubin-Spitz (1987) compared nine hearing aid conditions that varied systematically in the low and high frequencies by 6 dB/octave. Performance measures included paired-comparison judgments of intelligibility and speech discrimination scores on a nonsense syllable test. Results revealed that discrimination scores for each of 10 hearing-impaired subjects did not differ significantly across conditions, and rank correlations between the two measures ranged from -.49 to .76. Because subjects were permitted to adjust the output level for each condition, Levitt et al. (1987) speculated that this adjustment may have reduced the perceptual differences among conditions.

In a comparative study of prescriptive fitting procedures, Sullivan et al. (1988) ranked four prescriptive procedures using paired-comparison and speech recognition measures evaluated at three output levels. Post hoc analyses of significant differences were similar for the two test measures. At two of the three output levels, the differences in frequency-gain response between the four prescriptions were large enough to show significant differences in the speech recognition score.

A weakness of the paired-comparison technique concerns the trade-off between speech quality and intelligibility. For example, Thompson and Lassman (1969, 1970) found that a flat frequency response was preferred by listeners with high-frequency hearing loss despite findings that speech discrimination scores improved with high-frequency emphasis.

In support of the above findings, Punch and colleagues confirmed that listeners exhibited a strong preference for low-frequency amplification, and that quality judgments of continuous discourse were not related to judgments of intelligibility or nonsense syllable recognition (Punch & Beck, 1980; Punch et al., 1980; Punch & Parker, 1981). More recently, however, Punch and Beck (1986) provided evidence to suggest that low-frequency amplification may actually enhance syllable recognition in listeners with gradually sloping losses. Gordon-Salant (1984) similarly found that low-frequency amplification improved nonsense syllable recognition in listeners with flat losses, but degraded performance in listeners with steeply sloping, high-frequency losses.

The starting point in an adaptive paired-comparison method is usually specified by the best available prescriptive approach. It has been shown that the final result will either be as good if not better, certainly no worse than the best prescriptive approach. For example, Neuman et al. (1987) compared word recognition scores between the initial estimate of the optimum frequency-gain response (determined from a standard prescription) and the final condition selected by the simplex method. Seven of 8 subjects selected a frequency-gain response that differed from the initial estimate. Of those 7 subjects, 5 obtained higher speech recognition scores on the final response compared with the initial one, although only two of those scores were significantly higher.

From a more clinical perspective, Byrne and Cotton (1988) employed a paired-comparison test to evaluate performance with the revised National Acoustic Laboratories (NAL) prescription (Byrne & Dillon, 1986) following 2 to 3 weeks of experience with the new prescription. Subjects were instructed to judge pleasantness of speech in noise and intelligibility of speech in quiet. Three tournaments were conducted to compare the NAL prescription with conditions that varied systematically by 6 dB/octave in the low and high frequencies. Although NAL was determined to be more intelligible for 62 of 67 ears (44 subjects), a variation of NAL was determined to be more pleasant for 16 of the 67 ears. Byrne and Cotton (1988) concluded that a paired-comparison technique provided an efficient means of identifying patients who might prefer a variation of the original prescription.

Another advantage of the paired-comparison technique is flexibility. As noted earlier, there is no restriction in type of characteristic, stimulus, or dimension of sound quality that can be tested. For example, Levitt, Neuman, and Toraskar (1987) used a paired-comparison technique to select the optimal combination of coefficients for each of 8 subjects tested on a generalized form of compression. In a different application, Harris and Goldstein (1979) employed a paired-comparison procedure to evaluate hearing aid quality in quiet and reverberant environments. As mentioned previously, Punch (1978) used music stimuli as well as male and female speech to evaluate hearing aid quality.

To conclude, paired-comparison judgments of intelligibility and measures of speech recognition correspond best when the differences among aids are large enough to produce substantial differences in recognition scores. However, when the range among speech recognition scores is narrow, the paired-comparison technique may still show significant differences. For the former case, the paired-comparison technique is considered superior because of better reliability and efficiency. For the latter case, the paired-comparison technique provides information that is not available from speech recognition tests.

Given the advantages of the paired-comparison technique, its use in children is worth considering. However, there is no reason to suspect that this technique would be exempt from developmental constraints. Thus, in considering its use, the linguistic and cognitive factors required in making comparative judgments must first be examined.

Children's Comparative Judgments

Linguistically, comparisons form a class of inflectional morphemes composed of comparatives and superlatives. Comparatives reflect comparison between two items, and are formed by attaching the -_er_ suffix to an adjective (e.g., "A is tall_er_ than B."). Superlatives represent comparisons among three items or more, and are formed by adding -_est_ to the adjective (e.g., "A is the tall_est_ of the three.").

Children's knowledge of comparatives and superlatives is apparent between the ages of 4 and 7 years (Nicolosi, Harryman, & Kresheck, 1978; Streng, Kretschmer, & Kretschmer, 1978; Wood, 1976). In this development, deaf children lag behind normal-hearing children by 2 to 6 years and even longer (Cooper, 1967; Raffin, Davis, & Gilman, 1978). Comprehension of comparisons may be present in normal-hearing children as young as 2 1/2 years, and is generally acquired by the age of 4 1/2 years (Layton & Stick, 1979). Expressive use of comparisons follows.

According to Clark (1969), inflectional morphemes represent a surface transformation derived from a deeper structure comparing two or more base strings (e.g., "A is tall." "B is tall." "A is <u>more</u> tall than B."). Children between the ages of 3 and 4 years understand the concept of "more", whereas "less" is understood between the ages of 4 and 5 years (Donaldson & Balfour, 1968; Donaldson & Wales, 1970; Wannemacher & Ryan, 1978).

Semantic development also contributes to the acquisition of comparisons. The principle of lexical markedness asserts that polar adjectives (e.g., "tall-short") are acquired in stages (Clark, 1973; Clark, 1969). The unmarked, or positive end pole, is acquired first (Donaldson & Wales, 1970; Ehri, 1976; Layton & Stick, 1979; Wannemacher & Ryan, 1978). The unmarked adjective names the dimension, thus becoming neutralized. For example, the word "tall" reflects the dimension of height in "The man is six feet tall." In contrast, negative or marked members of the pair imply relational comparison such as "The man is short."

Inflectional morphemes are first attached to adjectives conveying size or having visual or tactile attributes. Gitterman and Johnston (1983) found that children between the ages of 4.5 and 8 years performed better on tasks that required production of comparisons having visual attributes (e.g., "tall/short") or tactile attributes (e.g., "soft/hard") than on tasks that combined the two (e.g., "wet/dry"). In addition, Nelson and Benedict (1974) found that children between the ages of 4 and 6 years demonstrated better comprehension of the comparative -<u>er</u> when associated with adjectives that were relative in nature (e.g., "tall" and "fat"), as opposed to adjectives defined as categorical (e.g., "round") or contrastive (e.g., "clean").

Use of adjectives was also found to be an important factor in testing the auditory abilities of children. For example, Andrews and Madeira (1977) discovered that children between the ages of 6 and 8 years were able to perform a pitch discrimination task using a motor response, but could not perform the task using verbal responses of "high-low" or "higher-lower." However, children could apply such terms to tasks involving spatial relations.

The cognitive ability associated with linguistic comparisons is seriation, the ability to order items on some dimension and to make transitive inferences about the ordinal relations (Inhelder & Piaget, 1964; Piaget, 1967). A standard seriation task would require the child to order a series of sticks differing in length and to make logical inferences about the transivity of different lengths (e.g., if A>B and B>C then A>C).

Children under 7 years of age typically experience difficulty performing standard seriation tasks (Inhelder & Piaget, 1964; Piaget, 1967). Very young children, however, indicate some knowledge of serial order. Children as young as 18 months have been shown to stack toy bricks of decreasing size (Inhelder & Piaget, 1964). Additionally, young children can differentiate between two sticks that differ in length and can recognize that one is longer than the other (Piaget, 1967).

Seriation ability has been associated with dimensionalization. Dimensionalization was defined by Miller (1979) as "the process that provides the child with information about properties that can be ordered dimensionally" (p. 97), such as shape, color, and size. In the early stages of this development, young children show the ability to extract the distinctive feature when forced to make a two-choice judgment. However, the ability to make a relative judgment from a dimension series is conceptually more difficult for the young child, and it develops in stages over a period of time.

In the experiments conducted by Piaget, children were asked to order a set of 10 sticks that varied in length. It was found that children 4 to 5 years of age were not able to order all 10 sticks, but could order a few from the set. Children 5 to 6 years of age could order the complete set, but did so with difficulty and by trial-and-error. At 7 years of age, children ordered the series of 10 sticks with ease and did so by means of an overall plan. Further, they could place two separate orderings into one-to-one correspondence (e.g., matching sticks with dolls of differing heights), and could make transitive inferences regarding the lengths (Ginsburg & Opper, 1969; Inhelder & Piaget, 1964; Piaget, 1967).

Although the ability to order length or size matures around the age of 7 years, children are not able to order a weight series until about the age of 9 years. The ability to seriate volume does not develop until 11 or 12 years of age (Piaget, 1967).

Furth (1964) has claimed that deaf children attain similar levels of performance on some cognitive tasks compared to normal-hearing children, particularly tasks that do not require language. He named seriation as one and reported its emergence in children between the ages 5 and 8 years.

In recent years, researchers have disputed the developmental milestones set by Piaget, noting that test procedures used in early studies did not adequately assess children's cognitive abilities. Bryant and Trabasso (1971) demonstrated that children 4 to 6 years of age were able to perform a transitive inference task after being trained first to discriminate among pairs from a set of five sticks differing in length and color. The training involved associating length with color, presenting pairs in reversible order and requiring the children to indicate which was "taller" or "shorter" ("big" or "little" for the 4-year-olds), and providing visual feedback.

A follow-up study conducted by Riley and Trabasso (1974) confirmed that use of both comparative terms ("taller" and "shorter") contributed to the conceptual learning of serial order, and was essential to successful performance on the transitive inference task. However, even after three or four training/testing sessions, only some of the children could correctly order all of the sticks.

Both studies also showed that pairs composed of the longest or shortest sticks (end-anchor pairs) yielded the highest proportion of correct responses, whereas the pairs composed of middle-length sticks generated a lower proportion of correct responses. These findings suggest that an absolute judgment may be easier to make than a relative one for children of this age.

The relationship between linguistic comparisons and seriation was investigated in several studies. In one, Ehri (1976) tested 40 children between the ages of 4 and 8 years on tasks of seriation and comprehension and production of adjectives. Seriation ability was shown to correspond with comprehension of comparative structures. That is, nonseriators (children 4 to 5 years of age) demonstrated less comprehension of comparatives than did young seriators (4 to 5 years of age) and older seriators (6 to 8 years of age).

The correspondence between production of comparatives and seriation ability was less clear. Ehri (1976) found that some nonseriators produced comparative structures but did not always comprehend the meaning of the comparison. In a later study, Gitterman and Johnston (1983) confirmed that children as young as 4.5 years of age produced the -<u>er</u> suffix, but did so in a noncomparative context. Children 7 years of age, however, successfully performed the seriation task while demonstrating receptive and expressive knowledge of comparative structures.

To conclude, the literature suggests that children's ability to make comparative judgments matures between the ages of 4 and 7 years. Clinically, optometrists have used a paired-comparison technique to help select corrective lenses for children 5 years of age and occasionally younger. They report that children of this age are able to choose the "better" of two visual images that differ in degree of blurring (personal communication: T. Feigenbaum, O.D., November, 1988; J. Garbus, O.D., October, 1988; D. Kirschen, O.D., Ph.D., November, 1988; V. White, O.D., December, 1988). It would be of interest to know whether children of the same age can make similar judgments with respect to auditory clarity.

Statement of Need

Hearing aid selection techniques appropriate for children are limited by constraints in testing. As hearing aids improve in technological sophistication and provide a wider range of options to the clinician, children may be denied benefit of these technological advances unless improved methods of hearing aid selection are developed for this age group. One technique, paired-comparison judgments, has yielded good results with adults and appears promising for children.

Although there is evidence to support the use of visual paired-comparison tests with children 5 years of age and older, the age at which children are able to make auditory paired-comparison judgments still needs to be determined. Furthermore, successful performance by normal-hearing children may or may not ensure successful performance by hearing-impaired children of the same age. Finally, it is unknown whether hearing-impaired children would be able to select appropriate hearing aids by a paired-comparison technique. A study was designed to explore these specific issues.

The Present Study

Three experiments were conducted to investigate use of the paired-comparison technique with children. The first two experiments determined the age at which normal-hearing and hearing-impaired children were able to make paired-comparison judgments of auditory clarity. The third experiment investigated the effectiveness of the paired-comparison technique to select appropriate hearing aids for hearing-impaired children. The specific aims of the study were as follows:

(1) To determine the age at which at least 50% of normal-hearing children can be expected to meet criterion on a paired-comparison task involving judgments of auditory clarity.

(2) To determine the age at which at least 50% of children with mild to moderately-severe sensorineural hearing loss can be expected to meet criterion on a paired-comparison task involving judgments of auditory clarity.

(3) To investigate whether a paired-comparison technique can be used effectively to select an appropriate hearing aid for hearing-impaired children who demonstrate the ability to perform an auditory paired-comparison task.

In the first two experiments, normal-hearing and hearing-impaired children of different ages were trained to judge auditory clarity of continuous discourse. Clarity, a perceptual dimension of sound quality, has been associated with intelligibility. A broadband, flat frequency response has resulted in judgments of clarity for normal-hearing listeners (Gabrielsson & Sjogren, 1979). Clarity judgments have also been associated with an increased high-frequency response and either a flat or decreased low-frequency response for hearing-impaired listeners (Gabrielsson, Schenkman, & Hagerman, 1988). In a study conducted with hearing-impaired adults, Tecca and Goldstein (1984) found a close correspondence between intelligibility and quality when subjects were instructed to base intelligibility ratings on clarity of speech.

In the present study, children were taught the concept of clarity through visually-blurred examples. A similar technique had been adopted by Studebaker and Sherbecoe (1988). They used visually-blurred photographs to train normal-hearing adults to judge speech intelligibility and quality by magnitude estimation. For the present study, the video portion of a videotape was altered by low-pass filtering. A videotape was utilized because it more closely represented continuous discourse. It's use also helped to maintain the child's interest, as such viewing has been shown to foster children's attentiveness (Anderson, Alwitt, Lorch, & Levin, 1979; Anderson & Lorch, 1983).

While the videotaped story progressed, each child was instructed to select the clearer of two conditions alternating in visual clarity. The visual clarity task was followed by similar tasks that assessed auditory-visual and auditory-only clarity. The auditory component of the videotape was altered by low-pass filtering for normal-hearing children and by frequency-gain shaping for hearing-impaired children.

The progression from visual clarity to auditory clarity was based on the premise that one modality reinforces the other. There is evidence to suggest that this is true. It has been shown that children are more attentive to visual information than to auditory when watching television (Pezdek & Hartmann, 1983; Pezdek & Stevens, 1984). The visual component directs the child's attention to material that may not be completely understood, thereby facilitating comprehension of the auditory component.

The age at which normal-hearing children accurately performed the auditory paired-comparison task targeted an age range for testing hearing-impaired children on the same task. The hearing-impaired children who successfully completed the auditory paired-comparison task continued on to Experiment 3.

Experiment 3 investigated the effectiveness of the paired-comparison technique to select appropriate hearing aids for hearing-impaired children. Each of seven frequency-gain shapes were paired with every other shape to form a round-robin tournament. An auditory recording of a second children's story was used to minimize learning effects and maintain interest. The round-robin tournament was repeated and the results of each were rank ordered and analyzed by correlation to assess test-retest reliability. In addition, phoneme identification scores were obtained for each frequency-gain shape. The results from each tournament were combined and correlated with the phoneme identification scores to assess the relationship between judgments of clarity and measures of speech intelligibility.

A successful outcome from this study would be indicated by reliable hearing aid rankings between test and retest, and rankings that correlated with phoneme identification scores. Such findings would support use of a paired-comparison technique with children, and would justify future studies to test the validity of the technique to select any number of hearing aid parameters.

II. EXPERIMENT 1

Experiment 1 determined the age at which at least 50% of normal-hearing children can be expected to meet criterion on a paired-comparison task involving judgments of auditory clarity. A training/testing protocol was designed to assess children's ability to judge relative clarity in a videotape altered by visual, auditory-visual, and auditory low-pass filtering.

Method

Subjects

English-speaking children between the ages of 4-0 and 6-5 (years-months) were recruited for Experiment 1. Five children from each of the following age groups were sought: 4-0 to 4-5; 4-6 to 4-11; 5-0 to 5-5; 5-6 to 5-11; 6-0 to 6-5. To be accepted, the child was required to pass a hearing and vision screening and a paired-comparison photograph test.

Puretone audiometry was performed with a Saico clinical audiometer (SC8) and TDH-39 headphones (MX-41/AR ear cushions) in an Industrial Acoustics Company (IAC 404A) single-walled, sound-treated room to select children with air-conduction thresholds of 20 dB HL (re: ANSI S3.6-1969) or better for octave intervals between 500 and

4000 Hz (American Speech-Language-Hearing Association, 1985). Visual acuity was screened at 10 ft using the Distance E chart (Bernell Corp., South Bend, Indiana) to select children with binocular vision not poorer than 20/30 for children 5 years and older, and 20/40 for children under 5 years of age (Lippmann, 1971).

A paired-comparison photograph test was constructed from the DLM sequential picture cards (DLM Teaching Resources, Allen, Texas). Pictures from a six-card series were reproduced in clear and blurry color photographs (Appendix A). To be selected as a subject, the child had to identify the "better" (i.e., clearer) photograph in each of the six pairs.

The Test of Auditory Comprehension of Language (TACL) (Carrow, 1973) was administered to document receptive language. Developed for children between the ages of 2 and 7 years, the TACL was chosen because it assessed comprehension of select adjectives and inflectional morphemes. It was of interest to know whether the subjects demonstrated comprehension of specific adjectives (big, fast, little, soft, two, more) and inflectional morphemes (smaller, taller, fattest), items correctly understood by 75% of children between the ages of 3 and 4 years tested on the TACL. Language testing was conducted by a certified speech/language pathologist.

Twenty-six children between the ages of 4-0 and 6-5 were evaluated. The children were enlisted from families (*n* = 9) and friends (*n* = 7) of House Ear Institute (HEI) staff, from families of patients seen in the children's center at HEI (*n* = 2), and from two preschools (*n* = 3) and one Sunday school (*n* = 5).

Twenty-five of the 26 children were selected to participate in the experiment. One child failed the hearing screening and was replaced. The 25 selected subjects (10 males, 15 females) ranged in age from 4-0 to 6-5 with a mean age of 5-2. The receptive language age ranged from 5-2 to 6-11 with a mean of 6-2. A listing of subjects is provided in Table 1. Subjects are identified from youngest to oldest by age group (4.0, 4.5, 5.0, 5.5, 6.0) and age order within the group (A, B, C, D, E).

The 4.0-year age group included 2 males and 3 females between the ages of 4-0 and 4-5 with a mean age of 4-2. The language age ranged from 5-2 to 5-11 with a mean of 5-6.

The 4.5-year age group included 3 males and 2 females between the ages of 4-6 and 4-11 with a mean age of 4-9. The language age ranged from 5-2 to 6-8 with a mean of 5-10.

The 5.0-year age group included 3 males and 2 females between the ages of 5-0 and 5-1 with a mean age of

Table 1. Normal-Hearing Subjects: Sex and Age.

Subject		Sex	Age (Years-Months)	
ID No.	Order of Entry	M/F	Chronologic	Language
4.0A	15	M	4-0	5-4
4.0B	25	F	4-1	5-3
4.0C	24	M	4-3	5-2
4.0D	20	F	4-3	5-9
4.0E	19	F	4-5	5-11
4.5A	14	M	4-6	5-4
4.5B	9	M	4-9	6-1
4.5C	13	F	4-10	6-8
4.5D	18	M	4-11	5-2
4.5E	12	F	4-11	5-9
5.0A	21	M	5-0	6-11
5.0B	16	M	5-0	5-5
5.0C	10	F	5-0	6-4
5.0D	8	F	5-0	6-5
5.0E	23	M	5-1	5-7
5.5A	7	F	5-6	6-5
5.5B	5	F	5-6	6-9
5.5C	11	F	5-10	6-4
5.5D	6	M	5-11	6-9
5.5E	3	F	5-11	6-11
6.0A	22	F	6-1	6-3
6.0B	17	M	6-2	6-10
6.0C	4	F	6-2	6-11
6.0D	2	F	6-4	6-11
6.0E	1	F	6-5	6-11

5-0. The language age ranged from 5-5 to 6-11 with a mean of 6-2.

The 5.5-year age group included 1 male and 4 females between the ages of 5-6 and 5-11 with a mean age of 5-9. The language age ranged from 6-4 to 6-11 with a mean of 6-8.

The 6.0-year age group included 1 male and 4 females between the ages of 6-1 and 6-5 with a mean age of 6-3. The language age ranged from 6-3 to 6-11 with a mean of 6-9. Subjects 6.0D and 6.0E were pilot subjects used in testing the protocol. Three additional subjects (6.0A, 6.0B, and 6.0C) were enlisted to complete the group of 5.

All subjects demonstrated comprehension of the inflectional morphemes assessed on the TACL. All but 5 subjects demonstrated comprehension of the specified adjectives. Subjects 4.0A, 4.0C, and 4.5D did not understand "more," Subjects 4.0B and 4.0C did not understand "soft," and Subject 5.0C did not understand "fast."

Testing was conducted at the House Ear Institute. An informed consent form was issued to each child's legal representative explaining the risks, benefits, and terms of the study. Time involved in testing was compensated at a rate of $10 per hour. Rewards of stickers were provided during testing, and a small toy was issued at the end of the session.

Instrumentation

Instrumentation for Experiment 1 included a video cassette recorder (VCR) and monitor, a loudspeaker, and a computer-controlled digital master hearing aid (Nielsen, 1989). A block diagram is shown in Figure 1 and described below.

The auditory signal was routed from the audio output of a VHS video cassette recorder (Panasonic AG-2200) to a preamplifier (Shure M67). From there it was directed to an IBM/AT personal computer-based digital master hearing aid system that utilized two Ariel DSP-16 digital signal processing boards. The analog signal was fed to the Ariel boards by way of an anti-aliasing (AA) filter (Frequency Devices) and two 16-bit, analog-to-digital converters (ADC), where it was converted to digital form. The digitized signal was processed by TMS320C25 digital signal processing chips, converted to analog form by the 16-bit, digital-to-analog converters (DAC), and routed to the A and B positions of a programmable switch box. The output from the switch box passed through an anti-imaging (AI) filter (Frequency Devices) to a programmable attenuator (Wilsonics PATT), limiter (dbx 165A), and power amplifier

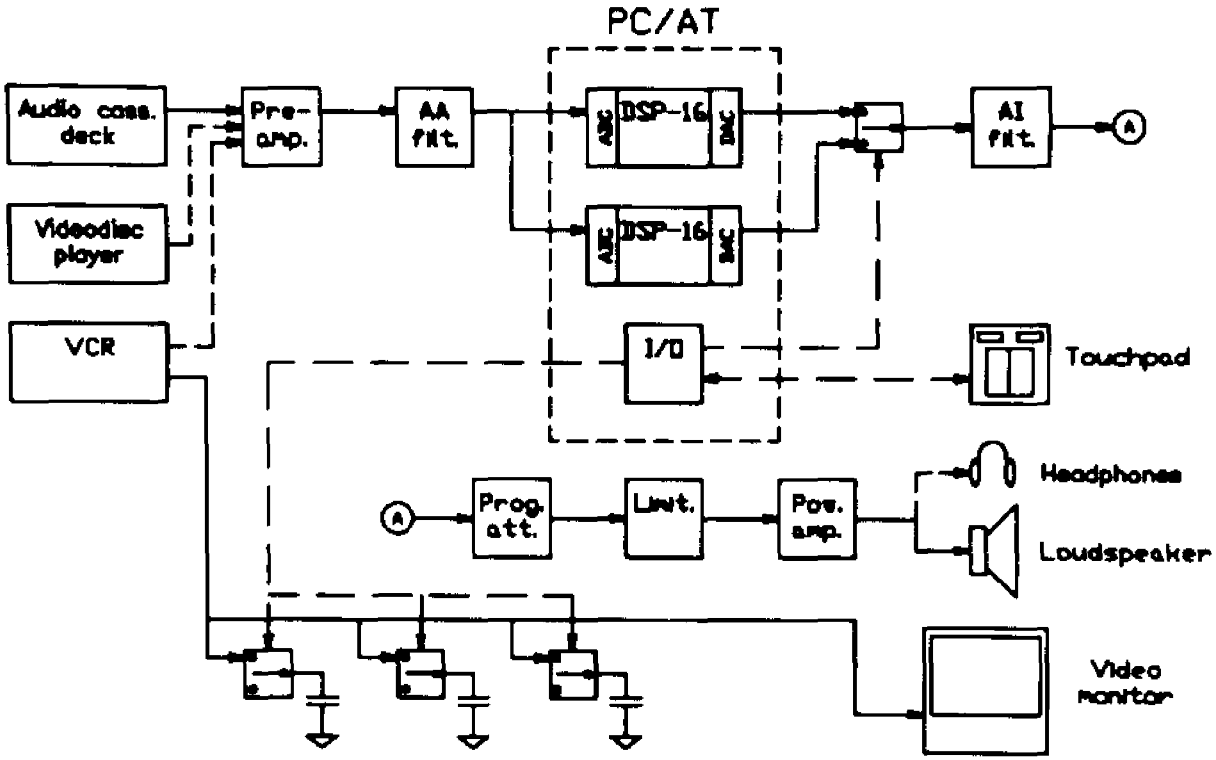


Figure 1. Block diagram of the test system used in Experiments 1, 2, and 3.

(Crown D-75) before being directed to the loudspeaker (JBL Studio Monitor 4406).

The video output was routed from the VCR to three switches of a programmable switch box. From there it was directed to the input of a 19-inch color video monitor (Sony PVM-1910). Three capacitors, connected to the output of the switches, were placed in parallel with the video monitor to create a set of low-pass filters.

A touchpad (KoalaPad) was connected to the personal computer via the game port of the input/output (I/O) port. The touchpad provided a means of switching between two conditions in making paired comparisons. Stickers were placed on the right side (moon sticker) and left side (star sticker) of a template to mark the areas to be touched when switching between conditions. The touchpad was also equipped with buttons for collecting the subject's responses.

Stimuli

The videotape used in the experiment was "Strega Nonna," an animated children's story written by Tomie dePaola (Prentice-Hall, Inc., 1975) and narrated by an adult female (produced by Patricia Greenfield, and borrowed from Kathy Pezdek, Ph.D., Claremont Graduate School, Claremont, California). The 8-minute story was appropriate for children between the ages of 3 and 8 years. The video component was presented in black and white to alleviate random color fluctuations that occurred as a result of the blurring process.

Four conditions of auditory clarity and four conditions of visual clarity were used in constructing the paired-comparison tests. Clarity ratings by 10 normal-hearing adults helped to establish the initial set of conditions.

Auditory low-pass filtering was implemented digitally by software using 253-tap finite impulse response (FIR) digital filters downloaded to the two Ariel boards. The auditory component of the videotape was low-pass filtered at 6000 Hz and digitized at a sampling rate of 20.16 kHz using the 16-bit, analog-to-digital converter. The digitized signal was processed, converted back to analog form via the 16-bit, digital-to-analog converter, and low-pass filtered at 6000 Hz by the anti-imaging filter.

The auditory low-pass filtered conditions initially selected (cut-off frequencies 4000, 2000, 1350, and 1000 Hz) were based on findings by Lawson and Chial (1982) on magnitude estimation and from an HEI pilot study on clarity rating. Subsequent testing with adult subjects revealed that adjacent conditions were difficult to differentiate (particularly 1350 vs. 1000 Hz) in the paired-comparison format. The final auditory conditions selected for the experiment were low-pass filtered with cut-off frequencies of 6000 Hz (condition A), 2000 Hz (condition B), 1000 Hz (condition C), and 500 Hz (condition D). With stopband ripple more than 60 dB down, filter slopes were greater than 90 dB/octave between passband and stopband frequencies.

The visual low-pass filtered conditions were also selected from pilot studies of clarity rating conducted with adults, and from informal paired-comparison tests conducted with adults and children. The four visual conditions were as follows: Broadband (condition A); 425 kHz (condition B); 142 kHz (condition C); and 85 kHz (condition D). Examples of the four conditions are shown in Figure 2.

Three paired-comparison tests were derived from the auditory and visual filtered conditions. The auditory test was composed of the four auditory low-pass filtered conditions and the visual test was composed of the four visual low-pass filtered conditions. The auditory-visual test combined the auditory with the visual conditions for simultaneous presentation. The four filtered conditions of each test were coupled in all possible combinations to

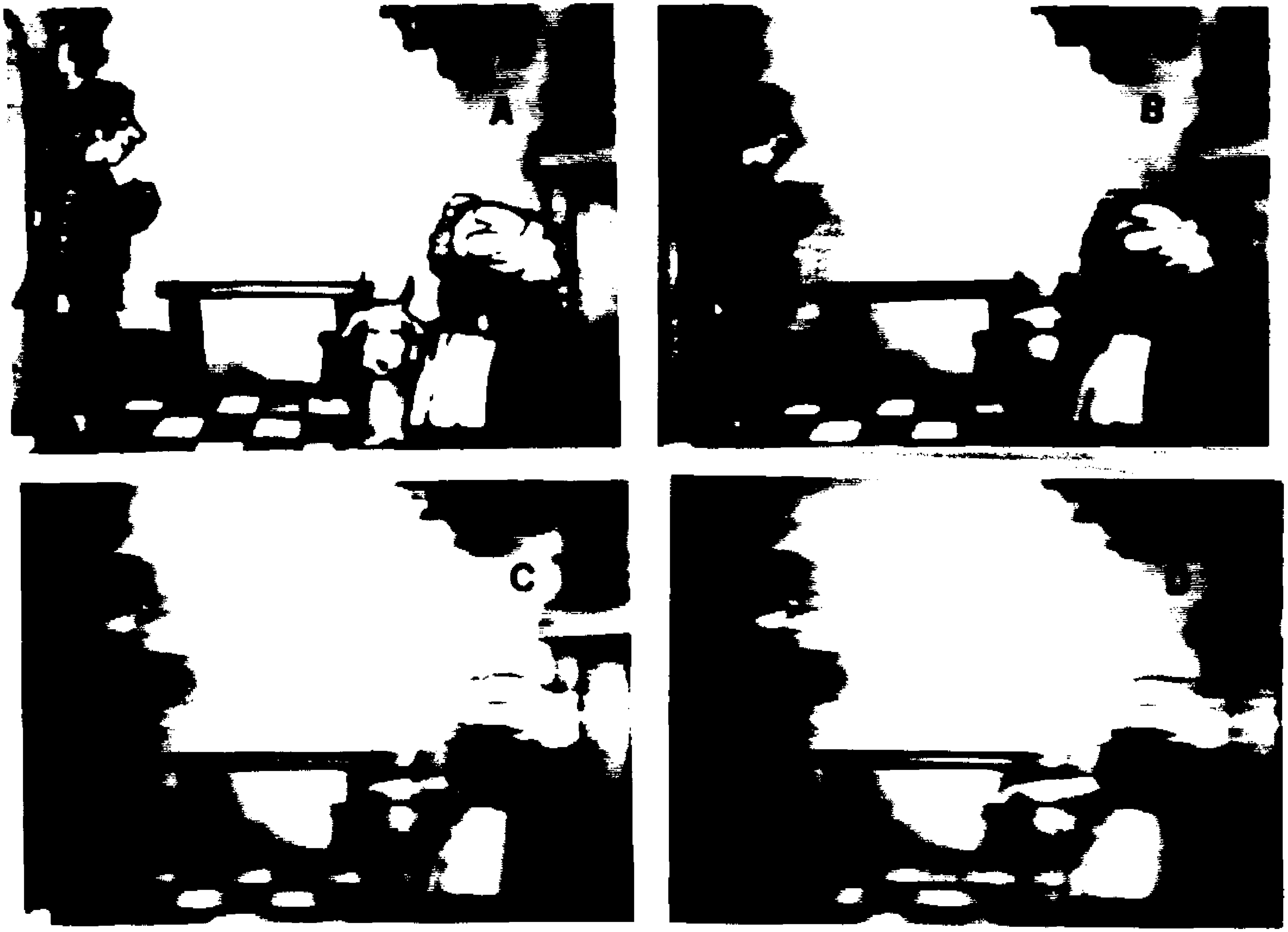


Figure 2. Four conditions of visual clarity: Broadband (condition A); 425 kHz (condition B); 142 kHz (condition C); 85 kHz (condition D).

form six pairs (AB, AC, AD, BC, BD, CD). Each pair was presented a second time, but in reversed order (BA, CA, DA, CB, DB, DC). The 12 pairs, presented in random order, comprised the paired-comparison test.

The training/testing protocol (see Procedure) was practiced on two 6-year-old children (Subjects 6.0D and 6.0E). Both children attained criterion levels of performance on all tests. Accordingly, the modified sets of conditions were accepted and incorporated into Experiment 1.

Procedure

Testing was conducted in an Industrial Acoustics Company (IAC 404A) single-walled, sound-treated room padded with 3-inch acoustic foam material. The subject was seated 1 meter from the video monitor and loudspeaker. The auditory component of the videotape was presented in the soundfield at a level of 70 dBA (integrated over a 100-second averaging time). The digital master hearing aid system was also stationed in the booth, thus making it possible for the tester to instruct the subject while typing commands into the computer. The noise level in the booth was 45.5 dBA with all instruments in operation.

The training/testing protocol assessed visual, auditory-visual, and auditory clarity, respectively.

Although the protocol was considered hierarchical, each level was attempted regardless of performance. However, if the subject failed to reach criterion on training, testing of that level was not attempted. If the subject did not reach criterion on the training component of the first level (visual clarity), the protocol was discontinued.

The paired-comparison task required the subject to select the "better" of two conditions while the videotape progressed. The subject was given practice operating the touchpad; first with clear and blurry photographs from the DLM cards, then with the videotape. Although the older subjects were able to operate the touchpad, the younger subjects were confused by it. Consequently, alternate techniques were implemented to train these subjects to perform the paired-comparison task.

The alternate techniques are described in the order they were attempted. For most subjects, the same technique was used throughout testing. For some subjects, however, the technique changed during testing (e.g., the subject took over operating the touchpad, or the subject required a change in task due to inattentiveness). The different techniques are described as follows:

(1) The subject controlled both the touchpad and response buttons to switch between conditions and record a response.

(2) The subject controlled the touchpad but gave a verbal response (e.g., "star" or "moon"). The tester pressed the response button to record the response.

(3) The tester controlled the touchpad while verbalizing the change (e.g., "This is the star. This is the moon. Which is better?"). The subject responded by pressing the response button to indicate the selected condition.

(4) The tester switched between conditions, verbally indicating the change. The subject verbalized the response (e.g., "star" or "moon").

(5) The tester switched between conditions and verbalized the change. The subject responded by putting tokens in appropriately labeled boxes.

The subject was allowed to switch between the two conditions as many times as needed. For those subjects who did not control the touchpad, the tester switched between conditions until a response was given.

During testing, some subjects responded impulsively before evaluating both stimulus conditions. In such instances, the tester reminded the subject to consider both conditions before making a decision. The pair was repeated and the response accepted if it was the same as the first. If different, the pair was again presented and this third response was accepted.

Training preceded testing for each step in the protocol. The training aspect presented a set of three pairs in order of difficulty, progressing from greatest to smallest contrast (AD, BD, BC). The subject's responses were hand-recorded by the tester and verbal feedback was provided. If the subject responded incorrectly, the pair was repeated. If the subject was again incorrect, the tester dropped down one step to the previous pair, or repeated that same pair if it was the easiest (AD). Training continued until six consecutive, correct responses were given for the three pairs presented twice (i.e., two sets). Training was discontinued if the subject failed to reach criterion after five sets were presented (or six sets if three correct responses were given on the fifth set).

Success in training permitted advancement to testing. In testing, responses were recorded by computer and feedback was not provided. The first step assessed visual clarity. The videotape was presented with sound, and only the visual component was altered in clarity. A minimum of

9 correct responses on the 12-item test indicated a passing score.

The next step contrasted auditory-visual clarity. The auditory component was altered simultaneously with the visual component. Additionally, the video picture was dimmed in brightness to increase the salience of the auditory component. Training preceded testing as described previously. A minimum of 9 correct responses indicated a passing score. If the subject did not reach criterion on training or testing, the level was repeated with brightness readjusted to normal.

The final step assessed auditory clarity. The video monitor was turned off. As before, training preceded testing. Following this test, a second auditory paired-comparison test was administered without training to assess repeatability. The youngest age group in which all 5 subjects obtained passing scores on the auditory test (9 or more correct responses) represented the age at which no less than 50% of children in the population could be expected to perform the task at the .05 probability level. That is, if 50% of children in the population can pass the test, then the probability of 5 out of 5 "passers" in the study sample occurring by chance is less than .05.

It is recognized that the use of a small sample size (i.e., 5 subjects per age group) increases the risk of making a Type II error. For example, if 90% of children in a population age group can pass the test, then the probability of 5 out of 5 subjects in the study passing is .59. The error probability is high (.41) that less than 5 subjects in this age group will pass the test. On the other hand, the probability of committing a Type I error is reduced because all 5 subjects in the sample are required to pass the test.

The total test time per subject ranged from 1 hr 30 min to 4 hr with a median duration of 2 hr 30 min. The selection protocol required 1 hr of testing. The paired-comparison protocol ranged from 30 min to 3 hr, with most subjects requiring between 1 hr and 1 hr 30 min of testing. Time involved in testing varied according to the number of tests presented, number of test sessions, amount of training required, and duration of breaks between tests. Each level of training and testing was approximately 10 min in duration. For 20 of the 25 subjects, all tests were completed in one session. Five subjects required a second session to complete testing.

Results

Paired-comparison test results revealed that performance was related to age. As groups increased in age, a greater number of subjects reached criterion on the four paired-comparison tests. These results are presented in Table 2 and Figure 3.

Inspection of the data reveals that no subject in the 4.0-year age group successfully completed visual clarity training and the protocol was discontinued for 4 of the 5 subjects. Subject 4.0E, the oldest child in the group, progressed to the auditory-visual and auditory levels even though failing visual clarity training. This child had an exceptionally long attention span, an impressive receptive language age (5-11), and a willingness to perform additional tasks. Although training at each level was unsuccessful, the A condition in the AD pair was always chosen as the clearer.

Two subjects in the 4.5-year age group (4.5C, 4.5E) obtained passing scores on all four tests. Subject 4.5B passed three tests but failed the second auditory test. Subject 4.5A passed the visual test but failed both auditory-visual tests (bright and dim), and did not reach criterion on auditory clarity training. Subject 4.5D failed the visual test and did not reach criterion on auditory-visual or auditory clarity training. These last two levels were assessed at a second session, thus ruling out fatigue as an influencing factor.

Three subjects in the 5.0-year age group (5.0A, 5.0B, 5.0C) reached criterion on the four paired-comparison

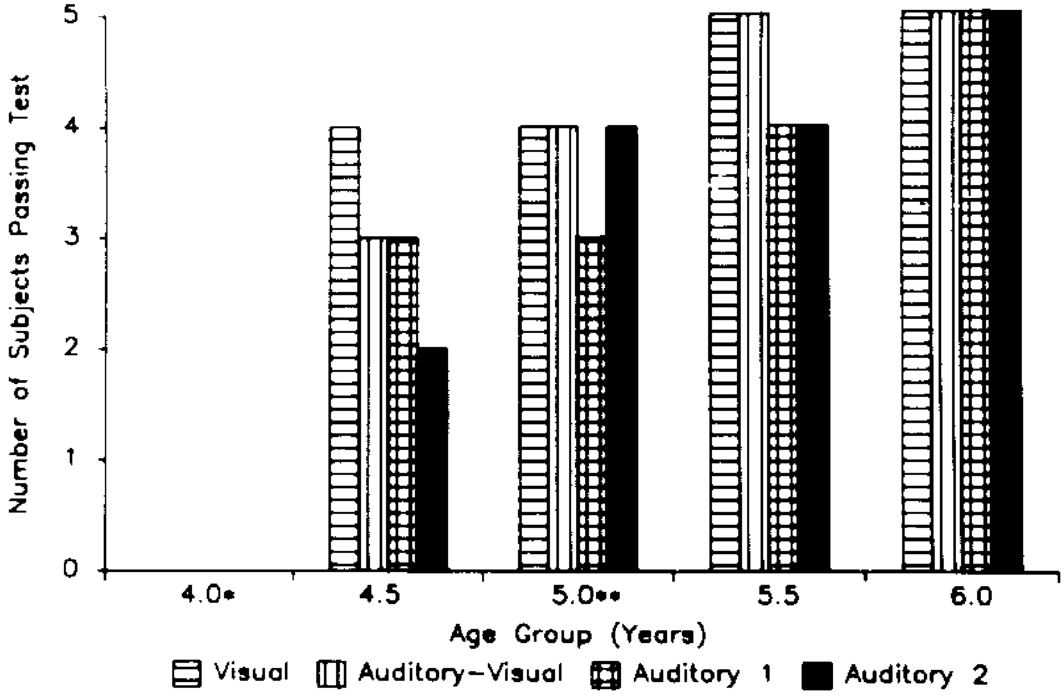
Table 2. Paired-Comparison Test Scores for Normal-Hearing Subjects.[a]

		Paired-Comparison Clarity Tests		
Subject	Visual	Auditory-Visual	Auditory 1	Auditory 2
4.0A	D/C	DNT	DNT	DNT
4.0B	D/C	DNT	DNT	DNT
4.0C	D/C	DNT	DNT	DNT
4.0D	D/C	DNT	DNT	DNT
4.0E	D/C	D/C	D/C	DNT
4.5A	9	7[b]	D/C	DNT
4.5B	9	9	11	8
4.5C	10	11	10	9
4.5D	8	D/C[b]	D/C	DNT
4.5E	9	11	9	9
5.0A	12	12	10	11
5.0B	10	10	11	10
5.0C	11	9	10	9
5.0D	D/C	DNT	DNT	DNT
5.0E	9	9	6	11
5.5A	10	11	10	9
5.5B	10	10[b]	8	5
5.5C	9	11	10	11
5.5D	12	11	9	12
5.5E	10	12	11	10
6.0A	12	11	12	12
6.0B	11	12	12	12
6.0C	11	10	11	10
6.0D	12	10	12	11
6.0E	11	10	11	12

[a] For the 12-item test, a score of 9 or better is passing.

[b] Supplemental auditory-visual test.

Note. D/C = Training discontinued. DNT = Did not test.



* 5 subjects did not pass visual training.

** 1 subject did not pass visual training.

Figure 3. Performance of normal-hearing subjects (N = 25) on the four paired-comparison clarity tests as a function of age group.

tests. Subject 5.0E passed three of the four tests but failed the first auditory test. Subject 5.0D did not reach criterion on visual clarity training (assessed at a second session) and the protocol was discontinued.

Four of 5 subjects in the 5.5-year age group reached criterion on the four tests. Subject 5.5B passed the visual and supplemental auditory-visual test, but failed the primary auditory-visual test (with brightness dimmed) and two auditory tests. The supplemental test and two auditory tests were performed during a second session.

All 5 subjects in the 6.0-year age group successfully achieved criterion levels of performance on the four paired-comparison tests. This finding suggests that at least 50% of children in this population age group could be expected to meet criterion on the auditory paired-comparison task at the .05 probability level.

The proportion of correct responses on individual test pairs (collapsed across tests) was calculated for age groups 4.5, 5.0, 5.5, and 6.0. These data are presented in Figure 4. It can be seen that the 6.0-year age group attained the highest proportion of correct responses followed by the 5.5- and 5.0-year age groups. The 4.5-year age group displayed the lowest proportion of correct responses.

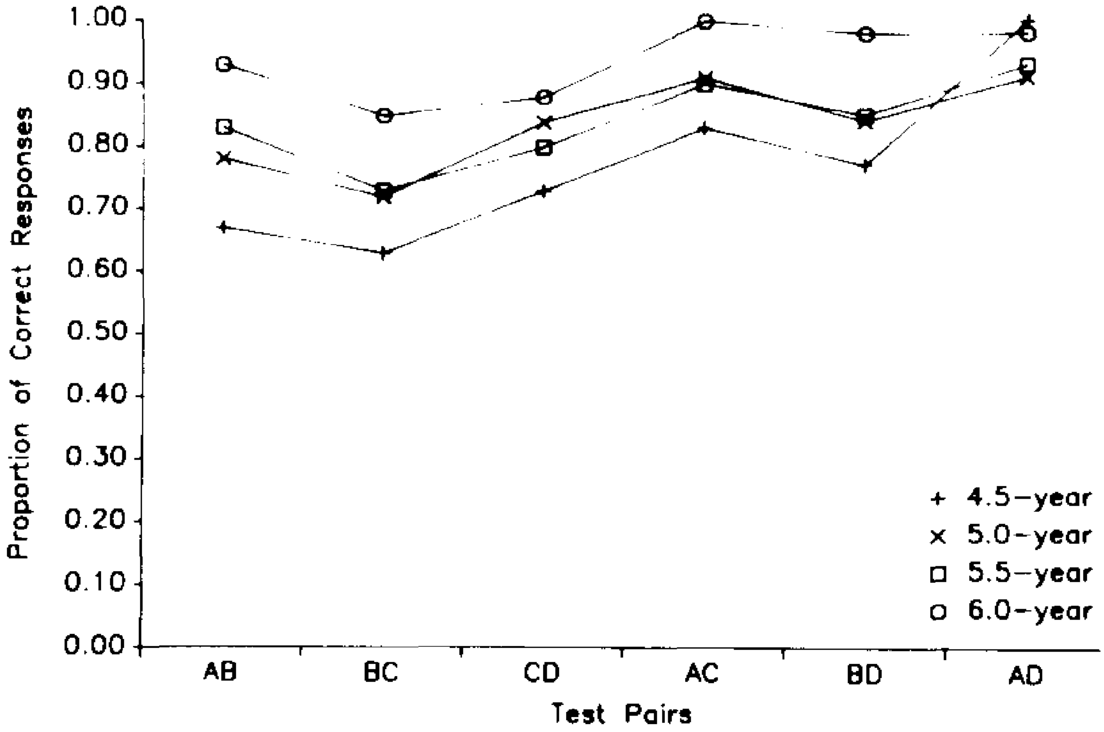


Figure 4. Proportion of correct responses on test pairs for four age groups of normal-hearing subjects combined over visual, auditory-visual, and auditory paired-comparison tests.

Despite age-group differences in performance, the response trends were similar. Subjects performed better on pairs with greater contrast (AD, AC, BD) than on pairs composed of adjacent conditions (AB, BC, CD). Further, there was a slight drop in performance on the BC and BD test pairs.

Collapsed across subjects, the proportion of correct responses on individual test pairs was calculated for the visual, auditory-visual, and auditory tests; results from the two auditory tests were pooled. These data are presented in Figure 5. It can be seen that response patterns were similar across pairs for the different tests except for the AB and BC pairs on the visual test. For those pairs, a reversal in pattern was evident.

Summary

Twenty-five normal-hearing children between the ages of 4-0 and 6-5 were tested on a paired-comparison task involving judgments of visual, auditory-visual, and auditory clarity. Results indicated that performance was related to age. With each increment in age group, a greater number of subjects successfully completed the tests. Although most subjects in the 4.5-year age group obtained passing scores on some or all tests, the age

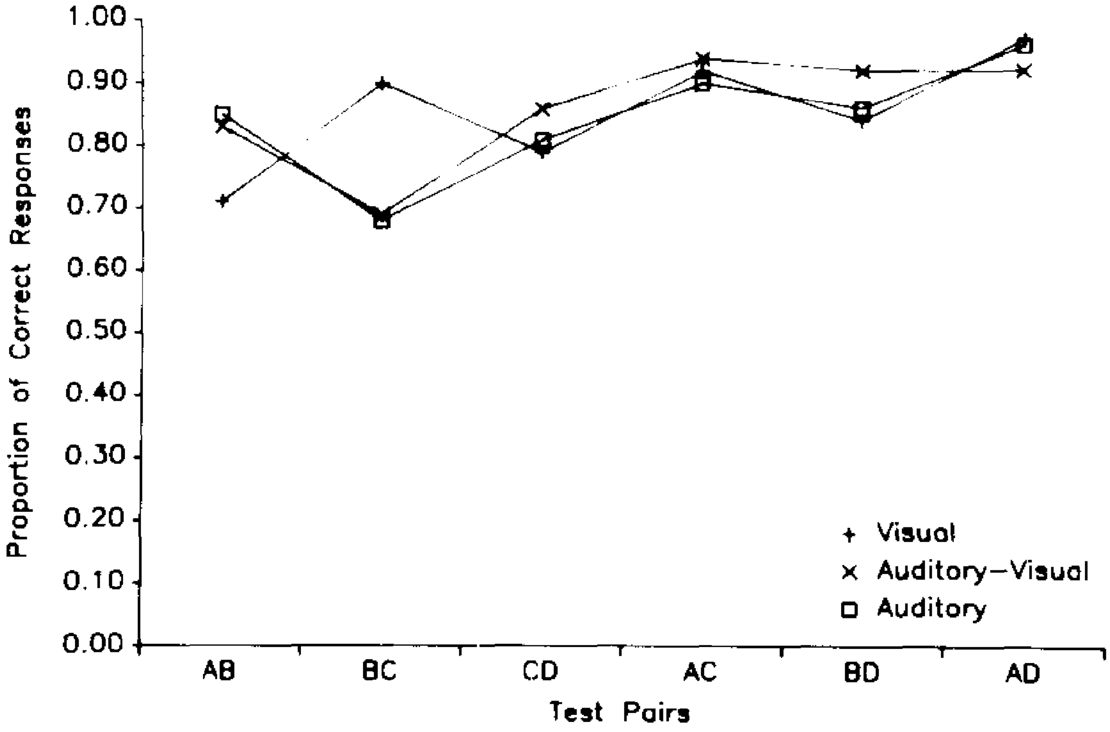


Figure 5. Proportion of correct responses on test pairs for the visual, auditory-visual, and auditory clarity tests for normal-hearing subjects.

group in which 5 out of 5 subjects reached criterion on the auditory paired-comparison test was 6.0 years.

The proportion of correct responses across test pairs was analyzed for the different age groups. The patterns were similar across test pairs despite age group differences in performance. The response patterns were also analyzed for the different tests. Again, the patterns were similar, although a discrepancy was noted between the AB and BC pairs when performance on the visual test was compared with performance on the auditory-visual and auditory tests.

III. EXPERIMENT 2

Experiment 2 determined the age at which at least 50% of children with mild to moderately-severe sensorineural hearing loss can be expected to meet criterion on a paired-comparison task involving judgments of auditory clarity. The training/testing protocol described in Experiment 1 was employed to assess hearing-impaired children's ability to judge relative clarity in a videotape altered by visual low-pass filtering and auditory frequency-gain shaping.

Method

Subjects

Based on findings from Experiment 1, hearing-impaired children of ages 5-6 (years-months) and older were recruited for Experiment 2. Children were required to speak English and to exhibit mild to moderately-severe, sensorineural hearing loss (26 to 70 dB HL, average threshold level re: ANSI S3.6-1969 for 500, 1000, and 2000 Hz), normal tympanograms (American Speech-Language-Hearing Association, 1990), and phoneme identification scores of 50% or better on the AB Word Lists (Boothroyd, 1968a). Hearing aid experience of 6 months or longer was also a selection requirement.

Additional selection criteria required that children pass the vision screening and paired-comparison photograph test described in Experiment 1. The TACL was administered to document receptive language age and to assess comprehension of specific morphemes and adjectives (see Experiment 1). Developed for normal-hearing children, the TACL is also appropriate for hearing-impaired children (Davis, 1977).

Eleven hearing-impaired children met the selection criteria. Ten children served as subjects and 1 child was used to pilot the procedures. The children were enlisted from House Ear Institute patient files (n = 9) and outside referral sources (n = 2).

The 10 subjects (5 males and 5 females) ranged in age from 5-7 to 7-2 with a mean age of 6-4. The receptive language age ranged from 5-2 to 6-10 with a mean of 6-4. The three-frequency average (3FA) for test ears ranged from 37 to 68 dB HL with a mean of 50.7 dB HL (SD = 11.6). Phoneme identification scores on the AB Word Lists (60 phonemes) ranged from 53% to 91% with a mean of 69.7% (SD = 14.27). The ear selected for testing was the aided ear (for monaural users) or the better-hearing ear (for binaural users). If thresholds were identical, then the test ear was selected at random.

A listing of subjects by chronologic and language age, educational status, hearing-loss history, puretone thresholds, phoneme identification scores, and hearing aid status is provided in Tables 3a, 3b, 3c, and 3d. Subjects are identified from youngest to oldest by age group (5.5 to 6.4 years, 6.5 to 7.4 years) and age order (A-J). A summary of means comparing the two age groups is provided in Table 4.

The 5.5 to 6.4-year age group included 3 males and 2 females between the ages of 5-7 and 6-4 with a mean age of 5-11. The language age ranged from 5-2 to 6-8 with a mean of 6-1. Three of the 5 subjects were in mainstreamed educational settings; Subject C was in preschool, Subject D in kindergarten, and Subject E in first grade. The 2 youngest subjects were in self-contained kindergarten classrooms; Subject A was in a class for slow learners, and Subject B was in a class for the hearing-impaired. All but Subject D received extra services that included speech therapy or academic tutoring.

Cause of hearing loss was attributed to birth trauma for Subjects B and E, and meningitis for Subject C (at age 2-6). Etiology and age at onset were unknown for Subjects

Table 3a. Hearing-Impaired Subjects: Sex, Age, and Educational Status.

Subject		Sex	Age (Years-Months)		Educational
Group	ID	M/F	Chronologic	Language	Status
5.5 to 6.4	A	M	5-7	5-2	Kindergarten
	B	F	5-8	6-3	Kindergarten
	C	M	5-9	6-7	Preschool
	D	M	6-2	5-9	Kindergarten
	E	F	6-4	6-8	First Grade
6.5 to 7.4	F	M	6-6	6-6	First Grade
	G	F	6-7	6-10	First Grade
	H	F	6-7	6-10	First Grade
	I	F	7-0	6-6	First Grade
	J	M	7-2	6-6	First Grade

Table 3b. Hearing-Impaired Subjects: Hearing-Loss History.

Subject		Age (Years-Months)		
Group	ID	Onset of Loss	Loss Diagnosed	Etiology
5.5 to 6.4	A	Unknown	3-3	Unknown, COM [a]
	B	Birth	4-6	Birth Trauma
	C	2-6	2-6	Meningitis
	D	Unknown	5-0	Unknown, COM
	E	Birth	3-6	Birth Trauma
6.5 to 7.4	F	Unknown	3-0	Unknown, COM
	G	Unknown	2-0	Unknown, COM
	H	Unknown	3-8	Unknown, COM
	I	Unknown	4-6	Heredity
	J	2-5	3-7	Meningitis

[a] COM = History of chronic otitis media.

Table 3c. Hearing-Impaired Subjects: Puretone Thresholds[a], Three-Frequency Average (3FA)[a], and Phoneme Identification Score.

Subject		Frequency (Hz)									Phoneme ID Score (%)
Group	ID	250	500	750	1000	1500	2000	3000	4000	3FA	AB Word Lists
5.5 to 6.4	A	70	70	60	65	70	70	70	65	68	53
	B	15	20	40	50	60	70	70	90	47	69
	C	25	35	45	60	55	20	25	25	38	88
	D	35	45	50	60	55	40	45	45	48	74
	E	30	30	40	45	55	60	55	45	45	66
6.5 to 7.4	F	40	55	60	65	70	70	75	75	63	54
	G	60	65	65	65	60	70	65	60	67	91
	H	15	20	15	20	50	70	65	70	37	61
	I	30	35	35	40	30	35	30	25	37	88
	J	55	60	55	55	60	55	55	65	57	53

[a] In dB HL re: ANSI S3.6-1969.

Table 3d. Hearing-Impaired Subjects: Hearing Aid Status.

Subject		Hearing Aid Status		
Group	ID	Ear Aided	Age Aided (Yrs-Mos)	Duration Use (Yrs-Mos)
5.5 to 6.4	A	Both	3-6	2-1
	B	Both	4-7	1-1
	C	Left	3-0	2-9
	D	Both	5-2	1-0
	E	Both	3-8	2-8
6.5 to 7.4	F	Both	3-9	2-9
	G	Both	2-0	4-7
	H	Left	4-0	2-7
	I	Both	4-9	2-3
	J	Both	6-0	1-2

Table 4. Summary of Means Comparing Two Age Groups of Hearing-Impaired Subjects.

	Age Group (Years)	
	5.5 to 6.4	6.5 to 7.4
Chronologic Age [a]	5-11	6-9
Language Age [a]	6-1	6-8
Age at Diagnosis [a]	3-9	3-4
Age HAs Fitted [a]	4-0	4-1
Duration HA Use [a]	1-11	2-8
3FA [b]	49.2 dB HL	52.2 dB HL
AB Word Lists	70.2%	69.4%

[a] In Years-Months.

[b] 3FA = Three-frequency average (dB HL re: ANSI S3.6-1969).

A and D. Both subjects also presented histories of chronic otitis media. Age at time of diagnosis for this group ranged from 2-6 to 5-0 with a mean of 3-9. Age when hearing aids were first fitted ranged from 3-0 to 4-7 with a mean of 4-0. Duration of hearing aid use (in years-months) ranged from 1-0 to 2-9 with a mean of 1-11.

Audiograms were bilaterally symmetrical for 4 of the 5 subjects. The 5th subject (C) exhibited a profound loss in the poorer-hearing, unaided ear. Puretone thresholds for test ears revealed configurations that were flat (Subjects A and D), gradually sloping (Subject E), steeply sloping (Subject B), and dropping only in the mid frequencies (Subject C). The 3FA ranged from 38 to 68 dB HL with a mean of 49.2 dB HL. Three of the 5 subjects exhibited air/bone gaps no greater than 10 dB at 250, 500, 1000, 2000, and 4000 Hz. Subjects C and D exhibited a 15-dB gap at 500 Hz. All 5 subjects displayed normal tympanograms. Phoneme identification scores ranged from 53% to 88% with a mean score of 70.2%.

The 6.5 to 7.4-year age group included 2 males and 3 females between the ages of 6-6 and 7-2 with a mean age of 6-9. The language age ranged from 6-6 to 6-10 with a mean of 6-8. All 5 subjects were mainstreamed in first-grade classrooms. All but Subject H received extra services including speech therapy or academic tutoring.

Cause of sensorineural hearing loss was unknown for 3 of the 5 subjects. These 3 subjects (F, G, and H) also presented histories of chronic otitis media. Hearing loss was attributed to hereditary factors for the 4th subject (I). The 5th subject (J) incurred a hearing loss at age 2-5 from meningitis. Age at time of diagnosis for this group ranged from 2-0 to 4-6 with a mean of 3-4. Age when hearing aids were first fitted ranged from 2-0 to 6-0 with a mean of 4-1. Duration of hearing aid use ranged from 1-2 to 4-7 with a mean of 2-8.

Audiograms were bilaterally symmetrical for 4 of the 5 subjects. The 5th subject (H) demonstrated an asymmetrical loss and used the hearing aid in the poorer-hearing ear. Puretone thresholds for test ears revealed configurations that were flat (Subjects G, I, and J), gradually sloping (Subject F), and steeply sloping (Subject H). The 3FA ranged from 37 to 67 dB HL with a mean of 52.2 dB HL. Air/bone gaps no greater than 10 dB were observed for 4 subjects. Subject G exhibited a 20 dB gap at 250 Hz. Tympanograms were normal for all 5 subjects. Phoneme identification scores ranged from 53% to 91% with a mean score of 69.4%.

All 10 subjects demonstrated comprehension of the inflectional morphemes assessed on the TACL. All but 4 subjects demonstrated comprehension of the specified adjectives. Subject D did not understand "big," and Subjects D, E, G, and J did not understand "fast." It was believed that the word "fast" may have been perceived as "fat" because all 4 subjects pointed to the picture of the elephant as opposed to the airplane--the correct picture.

As in Experiment 1, time involved in testing was compensated at a rate of $10 per hour. Rewards of stickers were provided during testing, and a small toy was issued at the end of each test session.

Instrumentation

Instrumentation was similar to that used in Experiment 1 with the following exceptions:

(1) For most tests, the signal was routed from the digital master hearing aid to TDH-49 headphones (model 51 ear cushions) or to the loudspeaker (JBL 4406). For some tests, however, the signal was routed from the audiometer to TDH-39 headphones (MX-41/AR ear cushions).

(2) The AB Word Lists were recorded on laserdisk. A laserdisk player (Pioneer Laservision Player LD-V4200) was used to route the signal to the audiometer or digital master hearing aid.

(3) The subjects' verbal responses on the AB Word Lists were recorded with a Marantz Stereo Cassette Recorder (PMD430) and Electro-Voice omnidirectional microphone (635A).

(4) The recorded passage of continuous discourse used in testing loudness discomfort level was delivered to the audiometer by a Vector Research Audio Cassette Deck (VCX-510).

(5) Tympanograms were obtained with the Grason-Stadler (GSI 28) tympanometer.

(6) Electroacoustic tests on subjects' personal hearing aids were performed with the Fonix 6500 Hearing Aid Test System and Fonix 6010 Sound Chamber (Frye Electronics, Inc.).

Stimuli

The videotaped children's story "Strega Nonna" (described in Experiment 1) was used to test the subjects' ability to make paired-comparison judgments. Four conditions of visual clarity and four conditions of auditory clarity were used in constructing the paired-comparison tests. The video component of the videotape was altered by low-pass filtering (visual conditions A, B, C, and D described in Experiment 1). The audio component of the videotape was altered by frequency-gain shaping.

Shaping was implemented instead of low-pass filtering because it was believed that hearing-impaired subjects would be less sensitive to differences among the low-pass filtered conditions (e.g., see Lawson & Chial, 1982). The frequency-gain shapes used in this experiment are described below.

Nine frequency-gain shapes, or digital filters, were custom designed for each subject. The filter shapes varied systematically in the low and high frequencies by 12 dB from a reference shape calculated from a standard prescription. The reference shape was calculated from the theoretically-based Byrne and Dillon (1986) NAL prescription.

Although 6-dB variations have been used in previous studies (e.g., Byrne & Cotton, 1988; Levitt, Sullivan, Neuman, & Rubin-Spitz, 1987), a 12-dB step was found to produce greater perceptible differences among frequency-gain shapes. Accordingly, the 12-dB step was selected for the present study.

The pivot point between low and high frequencies was 1000 Hz. The frequency of 1000 Hz was selected because it represented the center for octaves ranging from 250 to 4000 Hz. The upper boundary of the bandwidth was set to 4000 Hz to lessen chances of reaching the limiting level due to gain requirements at higher frequencies for some subjects.

A graphic representation of the nine frequency-gain shapes (re: NAL prescription) is presented in a 3 x 3 (nine-cell) matrix displayed in Figure 6. The center cell (HA#5) reflects the NAL prescription. Each row in the matrix shows a decrease (-12 dB), no change, or increase (+12 dB) in gain below 1000 Hz. Each column in the matrix shows a decrease (-12 dB), no change, or increase (+12 dB) in gain above 1000 Hz. The slope of change occurs between 500 and 1000 Hz for low frequencies, and between 1000 and 2000 Hz for high frequencies.

Frequency-gain shaping was implemented digitally by software using 128-tap FIR digital filters and a sampling frequency of 16.6667 kHz. The auditory signal was routed to the digital master hearing aid system where it was low-pass filtered at 5000 Hz by the anti-aliasing filter and digitized at a sampling rate of 16.6667 kHz by the 16-bit, analog-to-digital converter. The digitized signal was shaped by the frequency-gain filters downloaded to the two Ariel boards. The shaped signal was converted back to analog form by the 16-bit, digital-to-analog converter, and low-pass filtered at 5000 Hz by the anti-imaging filter.

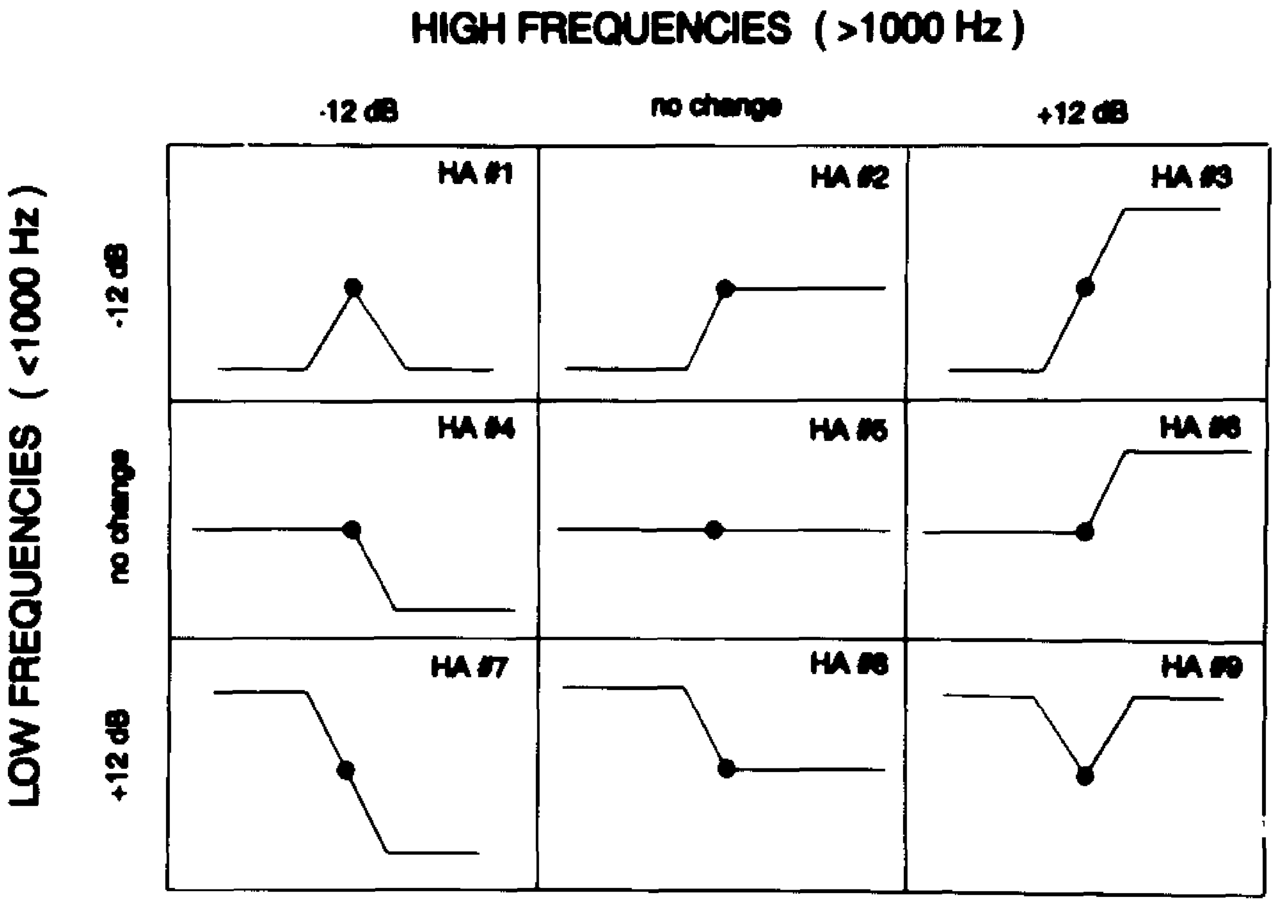


Figure 6. A 3 x 3 matrix representing nine frequency-gain shapes re: NAL prescription.

Because the NAL prescription calculated gain for a specified number of frequencies, a linear interpolation on a log-frequency scale was used to calculate 100 frequency samples (nearly evenly spaced) between 0 and 1/2 the sampling frequency to create smooth frequency response curves. An example of three filter shapes is shown in Figure 7. The figure presents shapes (HA#s) 1, 5, and 9 for Subject B. As can be seen, all three shapes pivot around 1000 Hz. HA#5 represents the NAL response for Subject B, whereas HA#s 1 and 9 depict the range in gain as a function of the 12-dB decrease and increase.

To verify that the nine filter shapes were accurate for each subject, output levels were measured at the headphone for specific frequencies. An input level of 60 dB SPL was used in performing these measurements. Results indicated that output levels were slightly under target, and differences were attributed to irregularities in the headphone response. Deviations from the target ranged from -0.1 dB in the mid frequencies to -2.5 dB at 4000 Hz. The output levels recorded for each subject are presented in Appendix B.

From the nine filter shapes, four needed to be selected for the paired-comparison test. In the original design of the study, the four shapes were to be determined individually from phoneme identification scores.

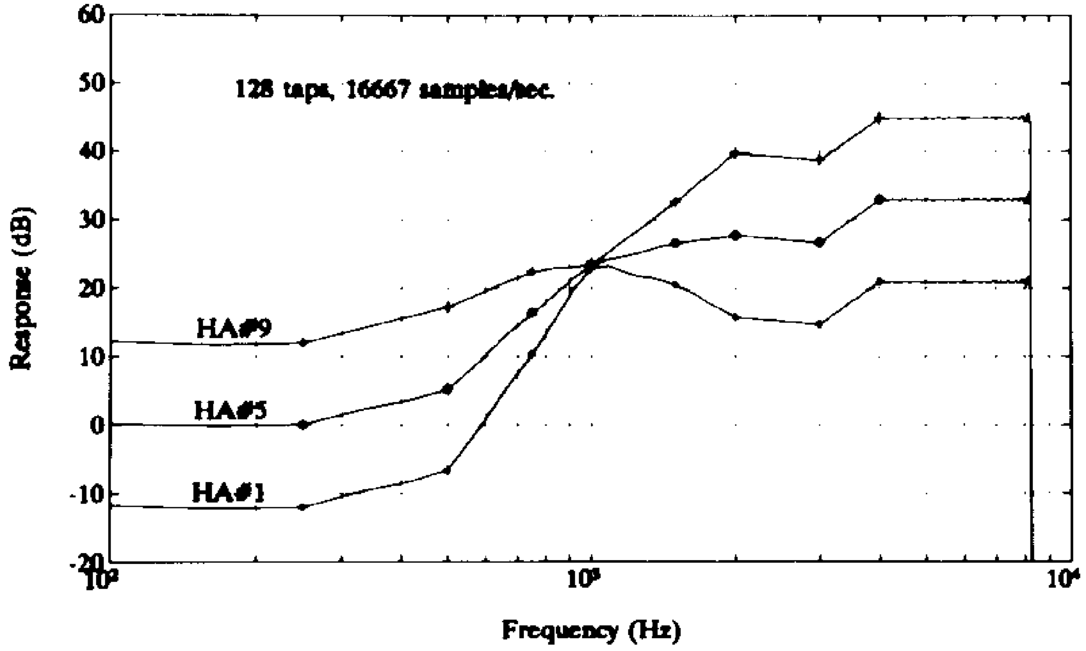


Figure 7. Frequency-gain shapes (HA#s) 1, 5, and 9 for Subject B.

Unfortunately, the range in test scores was too narrow to make this determination (see Appendix B for each subject's phoneme identification scores). Consequently, the same four shapes were selected for all subjects. Three of the shapes decreased systematically in the high frequencies (HA#6, HA#5, HA#4) and one shape decreased both in high and low frequencies (HA#1). It has been shown that such manipulations can influence perception of clarity for hearing-impaired listeners (Gabrielsson et al., 1988).

Preliminary tests conducted with Subject I and the pilot subject verified that most pairs formed by the four shapes were discriminable. Based on these findings, the four shapes were incorporated into the paired-comparison test as auditory conditions A (HA#6), B (HA#5), C (HA#4), and D (HA#1).

Phoneme identification was assessed with the AB Word Lists. These lists have been used with normal-hearing and hearing-impaired adults and children (Boothroyd, 1967, 1968a, 1970), and offer the following advantages: (1) Composed of 10 words, each list takes less testing time than other word recognition tests. (2) The words are administered in an open-set format and require only a verbal imitative response. (3) Phoneme scores are less affected by age than are word scores (Boothroyd, 1968b).

Recorded on laserdisk, the AB Word Lists were read by a female adult with general American dialect. Each of the 15 phonemically-balanced word lists was composed of 10 monosyllabic words (consonant-vowel-consonant). Each list was scored by calculating the number of correct phonemes, giving a maximum score of 30. For the present study, two lists were presented per condition, giving a maximum score of 60.

Loudness discomfort levels (LDLs) were obtained using a continuous discourse passage recorded on audio cassette tape. Recorded by a female adult, the discourse passage was the children's story, "Big Anthony and the Magic Ring," written by Tomie de Paola (Harcourt Brace Jovanovich, 1979). The speaker on the tape was the same speaker who recorded the AB Word Lists.

Procedure

Experiment 2 required three test sessions. The first session determined whether subjects met selection criteria. Test measures included air and bone puretone audiometry, tympanometry, the TACL (Carrow, 1973), the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1972), visual acuity screening, and the paired-comparison photograph test. In addition, coupler gain was measured for each subject's personal hearing aid(s) on the Fonix hearing aid test system.

The loudness discomfort level (LDL) for continuous discourse was obtained during the first session to determine the limiting level of the digital master hearing aid. The LDL method was adapted from Macpherson et al. (1989). Composed of four tasks, the first trained the subject to press a button to indicate "stop." The next two tasks trained the concept of "too much," and the fourth task used the previous steps to obtain LDL. Specific descriptions of the four tasks are detailed in Appendix C.

In testing LDL, the taped stimulus was routed from the audio cassette deck to the audiometer where it was delivered via TDH-39 headphones to the test ear. The subject was seated in the Industrial Acoustics Company (IAC 404A) single-walled, sound-treated room.

The taped stimulus was presented at 70 dB HL and increased in 10-dB increments. If the subject did not press the button by the time the level reached 100 dB HL (the limit of the audiometer), then testing was repeated. If there was no indication of discomfort after three ascending trials, then testing was discontinued and LDL was determined to be 100+ dB HL (118+ dB SPL). If the subject indicated that the signal was too loud at a level below 100 dB HL, then the signal was presented 5 dB below the level of response and increased in 5-dB increments. LDL was determined to be the level at which the subject responded at least three out of six times.

The visual paired-comparison test was also administered during the first session. Testing was conducted in the sound-treated room with the subject seated 1 meter from the video monitor and loudspeaker. The videotape was presented with sound, but only the video component was altered in clarity. The audio component was delivered through the loudspeaker at 70 dBA (integrated over 5 minutes). Personal hearing aids were adjusted to use setting.

The test protocol was similar to that described in Experiment 1. The subject was instructed to switch back and forth between two contrasting conditions and select the "better" of the two. Training preceded testing, and subjects were required to attain criterion levels on training to progress to testing. A score of 9 or better was required to pass the 12-item test.

The second test session assessed phoneme identification for the test ear in the following conditions: monaurally under headphones via the audiometer; monaurally under headphones for each filter shape; and in the soundfield with the personal hearing aid adjusted to use setting (see Appendix B for the recorded frequency-gain output levels). All tests were conducted with the subject seated in the sound-treated room.

For the first test condition, the laserdisk player was connected directly to the audiometer and the signal was delivered through TDH-39 headphones. The presentation level was determined for each subject during practice. Presentation levels ranged from 20 to 35 dB SL re: 3FA.

For the nine filter conditions, the signal was routed from the laserdisk player to the digital master hearing aid where it was processed and delivered through TDH-49 headphones at an input level of 60 dB SPL. The nine conditions were tested in random order.

For the personal hearing aid condition, the signal was routed to the digital master hearing aid where it was low-pass filtered at 6000 Hz and presented through the loudspeaker (1 meter from the subject) at 60 dB SPL. The nontest ear was plugged and muffed for this last condition.

Of the 15 word lists, 4 were used for practice. The practice lists were presented in random order, monaurally under headphones via the audiometer. The other 11 lists were presented twice in random order to equal 22 lists (2 lists per test condition).

No sound was allowed to exceed the subject's loudness discomfort level for speech (determined in the first session). The maximum power output of the digital master hearing aid was set to 120 dB SPL, the level recommended by Rintelmann and Bess (1988). This level was determined to be acceptable for Subjects A, C, D, G, I, and J. For Subject F, a level of 118 dB SPL was used, and for Subjects B, E, and H, a level of 103 dB SPL was used.

In testing phoneme identification, the subject was instructed to repeat each word while speaking into the microphone connected to the audio cassette recorder. The phonemes were scored by 2 listeners. The tester scored each subject's responses at the time of testing. A speech/language pathologist scored the responses from tape at a later time. Results from the Goldman-Fristoe Test of Articulation served as a reference for scoring. An analysis of variance was performed to test the main effect of scorer. Results indicated that differences were not statistically significant (see the Results section below). Accordingly, scores were averaged for reporting purposes.

The third session presented four paired-comparison clarity tests; auditory-visual (bright), auditory-visual (dim), auditory 1, and auditory 2. The auditory-visual tests combined the auditory and visual clarity conditions for simultaneous presentation. The video monitor was adjusted to normal brightness for the first auditory-visual test, then dimmed in brightness for the second. The two auditory tests were conducted with the video monitor turned off.

The subject was seated 1 meter from the video monitor with the audio portion routed through the digital master hearing aid to the headphones at an input level of 60 dB L_{eq}. This level differed from the 70-dB level used in Experiment 1, and was selected because there was less risk of the output reaching the limiting level. The limiter was set so that no sound exceeded LDL.

As described in Experiment 1, training preceded testing except for the second auditory test (used as a measure of repeatability). Although a score of 9 or better was required to pass each 12-item test, all tests were attempted regardless of performance on previous tests. However, if the subject failed to reach criterion on training, testing was not attempted.

The youngest age group in which 5 out of 5 subjects obtained passing scores on the auditory paired-comparison test represented the age at which no less than 50% of hearing-impaired children in the population age group could be expected to perform the task at the .05 probability level (See Experiment 1 for a discussion on this criterion).

Total test time per subject ranged from 3 hr 45 min to 5 hr 15 min with a mean duration of 4 hr 30 min. The first session required approximately 2 hr to complete the selection protocol. The second session required approximately 1 hr 30 min to present the AB Word Lists for 11 test conditions. The four paired-comparison tests administered in the third session were completed in about 1 hr.

Results

Paired-comparison test results for two age groups of hearing-impaired children showed that differences existed between and within groups. These results are presented in Table 5 and Figure 8.

Inspection of the data reveals that all subjects in the younger age group (5.5 to 6.4 years) passed the visual test but no subject passed all five of the tests. Subject A passed the visual and two auditory-visual tests, but failed to reach criterion on auditory clarity training and testing was not conducted. Subjects D and E failed the first (bright) and passed the second (dim) auditory-visual test, but failed both auditory tests. Subjects B and C passed one of the two auditory-visual tests and at least one of the two auditory tests.

Table 5. Paired-Comparison Test Scores for Hearing-Impaired Subjects.[a]

Subject		Paired-Comparison Clarity Tests				
Group	ID	Vis	AV-Bright	AV-Dim	Aud 1	Aud 2
5.5 to 6.4	A	12	12	12	D/C	DNT
	B	10	10	8	7	12
	C	10	7	11	9	9
	D	9	5	11	8	4
	E	10	8	12	4	8
6.5 to 7.4	F	11	11	11	10	8
	G	12	10	12	12	12
	H	11	10	11	10	9
	I	12	10	11	9	11
	J	10	12	12	9	10

[a] For the 12-item test, a score of 9 or better is passing.

Note: D/C = Training discontinued. DNT = Did not test.

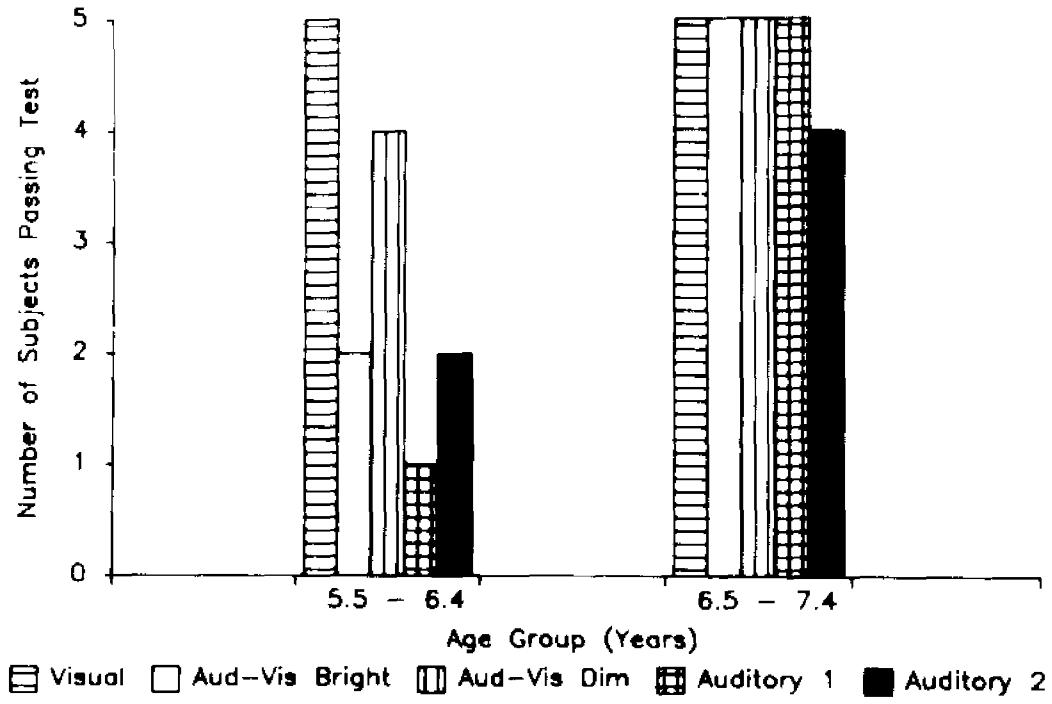


Figure 8. Performance of hearing-impaired subjects (N = 10) on the five paired-comparison clarity tests as a function of age group.

In contrast, all 5 subjects in the older age group (6.5 to 7.4 years) met criterion on the visual, auditory-visual, and auditory paired-comparison tests, although Subject F did not pass the second auditory test. These findings suggest that at least 50% hearing-impaired children in this population age group could be expected to pass the auditory paired-comparison test at the .05 probability level.

The proportion of correct responses on individual test pairs (collapsed across tests) was analyzed for the two groups. These data are displayed in Figure 9. It can be seen that a higher proportion of correct responses was obtained by the older group. The response pattern for the older group was relatively flat across test pairs, whereas the pattern for the younger group rose slightly.

Figure 10 shows the proportion of correct responses on individual test pairs for the visual, auditory-visual (bright and dim combined), and auditory (auditory 1 and 2 combined) tests collapsed across subjects. It can be seen that the highest proportion of correct responses was obtained for the visual test, closely followed by the auditory-visual test. The auditory test displayed the lowest proportion of correct responses. There was a slightly rising pattern across test pairs for the three tests.

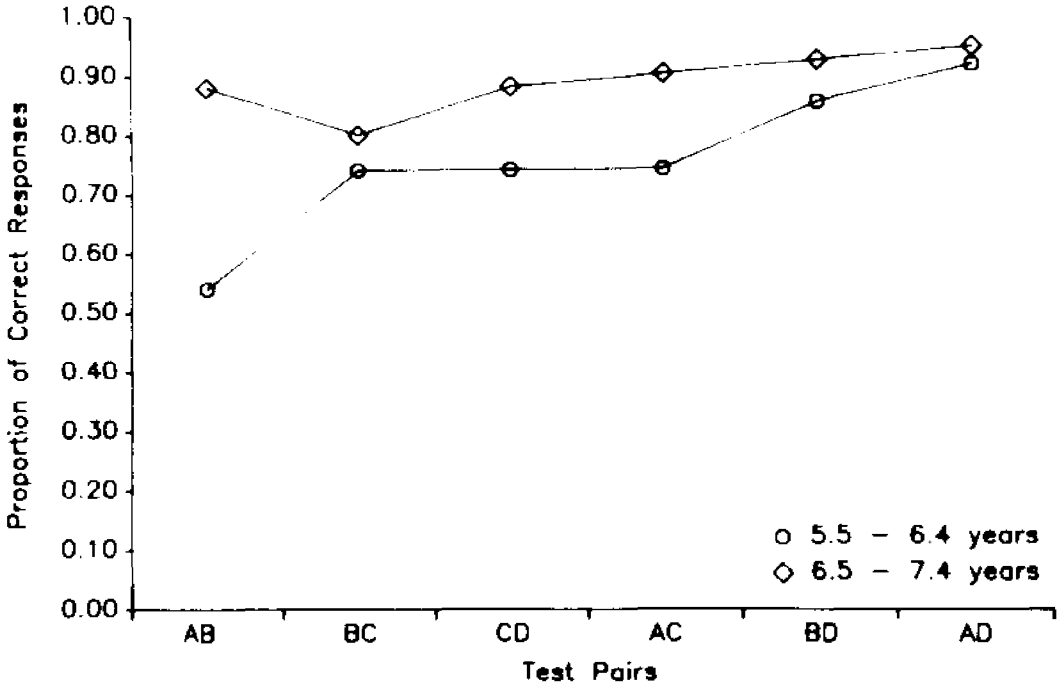


Figure 9. Proportion of correct responses on test pairs for two age groups of hearing-impaired subjects combined over the visual, auditory-visual, and auditory paired-comparison tests.

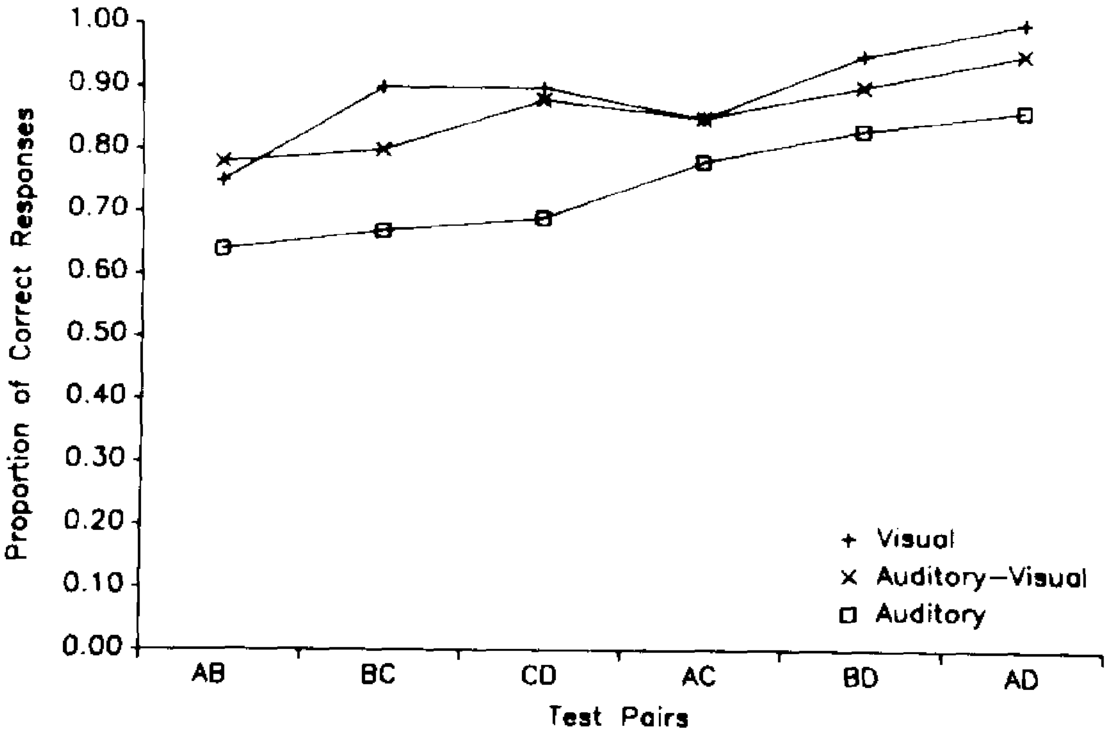


Figure 10. Proportion of correct responses on test pairs for the visual, auditory-visual, and auditory tests for hearing-impaired subjects.

To further analyze auditory abilities, phoneme identification scores for the nine filter shapes (hearing aids) were subjected to a three-way analysis of variance to assess the main effects of test (two levels), scorer (two levels), and hearing aid (nine levels). Only the main effect of hearing aid was found to be statistically significant, *F* (8, 144) = 24.87, *p* <.001. There were no significant interactions. The group means and standard deviations for each of the nine hearing aid conditions are presented in Table 6. Mean scores ranged from 25.3% for HA#1 to 65.1% for HA#6.

Post hoc analysis using the Tukey HSD test revealed that HA#1 differed significantly from HA#s 3, 5, 6, 7, 8, and 9, and HA#s 2, 4,and 7 differed significantly from HA#s 6 and 9 (*p* <.05). With respect to the auditory paired-comparison test, only paired conditions AC, BD, and AD were shown to be significantly different by phoneme identification scores.

Summary

Ten hearing-impaired children between the ages of 5-7 and 7-2 were tested on a paired-comparison task involving judgments of visual, auditory-visual, and auditory clarity. All 10 subjects passed the visual test. The 5 subjects in the younger age group (5.5 to 6.4 years)

Table 6. Mean Phoneme Identification Scores and Standard Deviations for 10 Hearing-Impaired Subjects on Nine Hearing Aid Conditions.

	Phoneme Identification Test (AB Word Lists)	
Hearing Aid	Mean Score (%)	Standard Deviation
1	25.3	14.1
2	42.3	23.1
3	55.5	26.1
4	40.5	17.8
5	51.1	18.4
6	65.1	22.8
7	44.3	15.2
8	55.0	20.4
9	64.6	18.3

passed one of two auditory-visual tests, and 2 subjects passed one or both of the two auditory tests. In contrast, all 5 subjects in the older age group (6.5 to 7.4 years) passed both auditory-visual tests and one or both of the two auditory tests.

Analyses of response patterns across test pairs reflected performance differences between age groups and between tests. The auditory paired-comparison test was found to be more difficult than either the visual or auditory-visual tests for the hearing-impaired subjects. Statistical analysis of phoneme identification scores revealed that some conditions were not significantly different.

IV. EXPERIMENT 3

Experiment 3 investigated whether a paired-comparison technique can be used effectively to select hearing aids for hearing-impaired children who demonstrate the ability to perform the auditory paired-comparison task. Seven hearing aid conditions competed in two paired-comparison, round-robin tournaments to assess test-retest reliability. The number of preferences for each hearing aid was correlated with phoneme identification scores to assess the relationship between judgments of clarity and measures of speech intelligibility.

Method

Subjects

Eight hearing-impaired children served as subjects in Experiment 3. Subjects included the 7 in Experiment 2 who successfully met criterion on one of the two auditory paired-comparison tests (Subjects B, C, F, G, H, I, and J), and a newly enlisted subject (K) recruited from House Ear Institute patient files.

Subject K (male) was 7-10 (years-months) and mainstreamed in a second grade classroom. Cause of hearing loss and age at onset were unknown, although there was a history of chronic otitis media. Sensorineural hearing loss was diagnosed at age 6-2, and hearing aids were fitted at age 7-3. Duration of hearing aid use was 0-7. Subject K exhibited a gradually sloping audiometric configuration with a three-frequency average (3FA) of 37 dB HL (re: ANSI-1969) for the test ear. The phoneme identification score on the AB Word Lists was 88% for the 60-item test.

On the five paired-comparison clarity tests described in Experiment 2, Subject K passed the visual, both auditory-visual, and the second of two auditory tests. Hearing aid output levels and phoneme identification scores have been included in Appendix B.

The 8 subjects (4 males and 4 females) participating in Experiment 3 ranged in age from 5-8 to 7-10 with a mean age of 6-8. The age when hearing loss was diagnosed ranged from 2-0 to 6-2 with a mean of 3-9. Hearing aids were first fitted at a mean of age of 4-5, and ranged from 2-0 to 7-3. Duration of hearing aid use ranged from 0-7 to 4-7 with a mean of 2-3. The mean 3FA for test ears was 47.9 dB HL (SD = 11.89), and the mean phoneme identification score on the AB Word Lists was 74% (SD = 15.46).

The time involved in testing was compensated at a rate of $10 per hour. Rewards of stickers were provided during testing, and a small toy was issued at the end of each test session.

Instrumentation

Instrumentation for Experiment 3 included the Marantz Stereo Cassette Recorder (PMD430) and digital master hearing aid system (described in Experiment 1). The auditory stimulus was routed from the digital master hearing aid to TDH-49 headphones (model 51 ear cushions).

Stimuli

Paired-comparison and same/different tests were conducted in Experiment 3. The auditory stimulus used for both tests was the audio cassette recording of the children's story, "Big Anthony and the Magic Ring," described in Experiment 2.

The paired-comparison test was constructed from seven of the nine frequency-gain shapes (hearing aids) used in the previous experiment. Each of the seven shapes was paired with every other shape to form a round-robin tournament. Seven shapes were utilized instead of nine to reduce the number of pairs from 36 to 21, thereby making the procedure more manageable for use with children. HA#s 1 and 9 were the shapes deleted because they represented the extreme gain conditions and were unlikely to be among the winning selections.

Each of the 21 pairs in the round-robin tournament was presented three times in random order to form a 63 paired-item test. The 63 pairs were presented a second time in random order to assess test-retest reliability. Combining the two paired-comparison tournaments resulted in 126 pairs, or six presentations of each pair. Order of presentation (e.g., AB and BA) was randomized across the six presentations.

The same/different test included only those pairs from the paired-comparison test in which there was no clear preference (i.e., if the same preference was indicated at least five out of the six presentations on the paired-comparison test, then that pair was not included in the same/different test). For each pair included, five same and five different pairs were formed to comprise a set (e.g., AB, BA, AA, BB, AB, BA, AA, BB, AB, AA). During testing, two sets of stimuli were interleaved at random.

Procedure

Experiment 3 required two test sessions. The two paired-comparison, round-robin tournaments were administered in the first session, and the same/different test was administered in the second. Subjects were tested in an Industrial Acoustics Company (IAC 404A) single-walled, sound-treated room. The auditory stimulus was delivered monaurally under headphones to the test ear at an input level of 60 dB L_{eq}.

The round-robin tournament was similar in structure to the auditory paired-comparison test described in Experiment 2. The subject was instructed to listen to the children's story presented on audio cassette tape. With use of the touchpad, the subject switched between two hearing aid conditions and indicated which of the two conditions sounded "better." The response was a judgment and, consequently, could not be scored as either correct or incorrect. No training or feedback were provided. A 15-minute break separated the two paired-comparison tournaments.

The same/different test also required use of the touchpad. While the children's story progressed, subjects alternated between two conditions and verbally indicated whether the two conditions sounded the same or different. Subjects were first trained on HA#s 2-8, 2-2, 8-2, and 8-8. The four pairs were presented until each was identified correctly by the subject. Feedback was provided during training. Testing followed with no feedback. For each set of 10 pairs (5 same and

5 different), a minimum of 8 correct responses determined that the two hearing aids being contrasted were perceived to be different.

Each session in Experiment 3 required approximately 1 hr. During the first session, each 63-item paired-comparison test required about 15 min to complete. The number of same/different sets presented in the second session varied from subject to subject.

Results

Test results from each paired-comparison, round-robin tournament were tabulated by number of preferences per hearing aid condition. The maximum number of times an aid could be selected in each test was 18. The two sets of results per subject were ranked from highest (1) to lowest (7) and correlated using the Spearman rank correlation coefficient (Siegel, 1956).

The number of preferences, hearing aid rankings, and correlations for each subject are shown in Table 7. The results showed that correlations between test and retest were strong (above .80) for Subjects B, G, H, J, and K; moderate (between .50 and .80) for Subject H; and weak (below .50) for Subjects C and F. It should be noted that

Table 7. Results on Two Paired-Comparison Tournaments: Number of Preferences (Ranking of Hearing Aids) and Spearman Rank Correlation Coefficients.

	Subject							
	B	C	F	G	H	I	J	K
Test 1								
HA #								
2	3 (7)	13 (1)	10 (2.5)	3 (6.5)	2 (7)	6 (6.5)	6 (6)	4 (6)
3	6 (6)	12 (2)	8 (5.5)	10 (4)	8 (5)	8 (4.5)	10 (4)	2 (7)
4	7 (5)	6 (7)	10 (2.5)	3 (6.5)	6 (6)	6 (6.5)	0 (7)	8 (4.5)
5	11 (3)	7 (5)	6 (7)	4 (5)	8 (3)	8 (4.5)	8 (5)	8 (4.5)
6	11 (3)	11 (3)	9 (4)	14 (2)	12 (4)	12 (2)	16 (1)	9 (3)
7	11 (3)	7 (5)	8 (5.5)	12 (3)	9 (2)	9 (3)	12 (2)	15 (2)
8	14 (1)	7 (5)	12 (1)	17 (1)	14 (1)	14 (1)	11 (3)	17 (1)
Test 2 (Retest)								
HA #								
2	2 (7)	7 (5.5)	6 (7)	2 (7)	3 (7)	4 (7)	7 (6)	3 (6.5)
3	10 (4.5)	12 (1.5)	9 (4)	7 (4)	8 (5)	11 (1.5)	9 (4)	3 (6.5)
4	6 (6)	10 (3.5)	7 (6)	5 (6)	6 (6)	8 (6)	4 (7)	8 (4.5)
5	10 (4.5)	12 (1.5)	9 (4)	6 (5)	9 (4)	10 (3.5)	8 (5)	8 (4.5)
6	11 (2.5)	7 (5.5)	9 (4)	14 (2)	11 (3)	9 (5)	10 (3)	13 (2.5)
7	11 (2.5)	5 (7)	10 (2)	13 (3)	12 (2)	10 (3.5)	11 (2)	13 (2.5)
8	13 (1)	10 (3.5)	13 (1)	16 (1)	14 (1)	11 (1.5)	14 (1)	15 (1)
Spearman Rank Correlation Coefficients								
r_s =	.89	-.10	-.06	.99	.96	.63	.86	.98

Subjects C and F were inattentive and required prompting throughout testing.

The average correlation between test and retest for the group of 8 subjects was .68. The mean number of preferences for each test are displayed graphically in Figure 11. The results show a close correspondence between tests across the seven hearing aid conditions.

The occurrence of ties was also noted for some subjects. Same/different tests were conducted to determine whether inconsistencies in judgments could be attributed to discrimination difficulties. As noted earlier, the hearing aids being contrasted formed a set of 10 pairs: 5 pairs were acoustically different (e.g., AB or BA); and 5 pairs were acoustically the same (e.g., AA or BB).

Unfortunately, the results were difficult to interpret because some subjects identified pairs that were acoustically the same as being "different" (possibly due to the varying nature of continuous discourse). Consequently, the two hearing aids being contrasted were often categorized as "not different" according to the criterion (8/10) established. Nevertheless, results are summarized below.

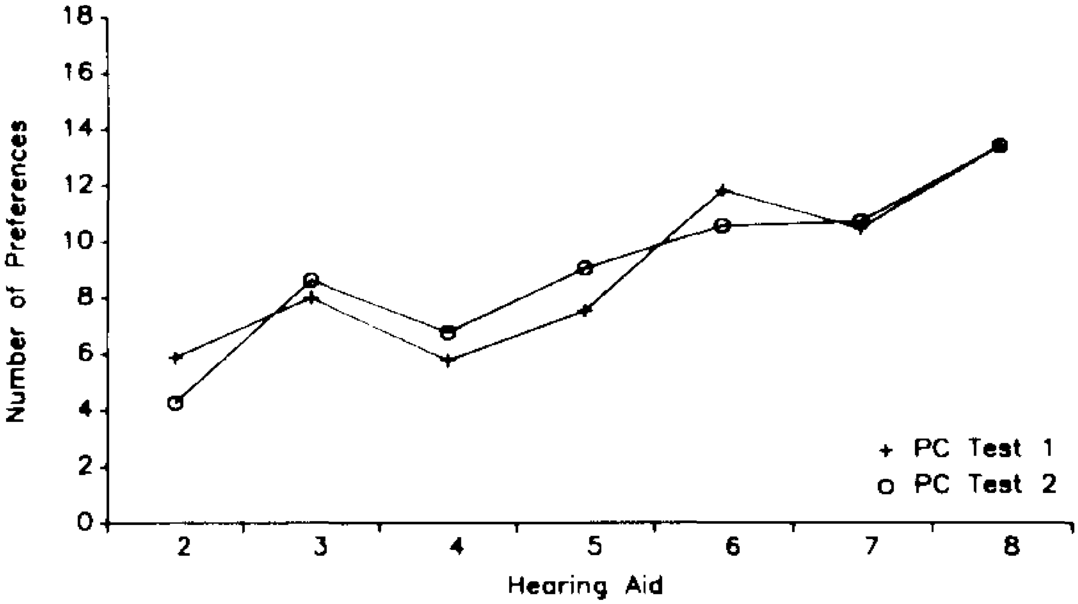


Figure 11. Mean number of preferences by hearing-impaired subjects ($N = 8$) on two paired-comparison, round-robin tournaments across hearing aids 2 through 8.

Of the 8 subjects tested, Subject C did not complete testing, and Subject I successfully met criterion on every set of 10 pairs tested. Of the remaining 6 subjects, 5 subjects did not differentiate between hearing aids 2-3; 4 subjects did not differentiate between hearing aids 3-4, 3-5, and 7-8; 3 subjects did not differentiate between hearing aids 2-5, 3-6, 5-6, and 6-7; and 2 subjects did not differentiate between hearing aids 2-4, 4-7, 5-8, and 6-8. Those hearing aids not differentiated by only 1 subject are not reported.

An alternate analysis was performed to determine whether differences among phoneme identification scores (AB Word Lists) were significantly different across the seven hearing aid conditions for a given subject. Each subject's mean score and variance (observed and expected) were calculated across the seven conditions. The expected variance was derived by assuming a binomial sampling distribution about the subject's mean score. Comparisons were evaluated using a chi-square distribution.

These data are presented in Table 8. It can be seen that the observed variance of phoneme identification scores was not significantly greater than the expected binomial variance for Subjects C, F, and J. This indicated that the effect of hearing aids was not significant for these 3 subjects. As noted earlier,

Table 8. Mean Phoneme Identification Scores and the Observed Variance versus Expected Variance across Hearing Aids 2 through 8 for Individual Subjects.

Subject ID	Phoneme ID Score Mean (%)	Variance	
		Observed	Expected
B	43.9	122.4 [a]	41.1
C	48.9	93.2	41.6
F	36.7	66.0	38.7
G	53.7	311.3 [b]	41.4
H	62.6	339.9 [b]	39.0
I	78.1	210.1 [b]	28.5
J	35.0	91.5	40.0
K	73.7	212.9 [b]	32.1

[a] Significantly greater than expected variance ($p < .01$).

[b] Significantly greater than expected variance ($p < .001$).

Subjects C and F also exhibited weak correlations between the two paired-comparison tournaments. For the remaining 5 subjects, the observed variance was significantly greater than the expected binomial variance ($p < .01$), indicating a significant effect of hearing aids. As was noted, these 5 subjects exhibited moderate to strong correlations between the two paired-comparison tournaments.

To examine the relationship between paired-comparison judgments and phoneme identification scores, the number of preferences per hearing aid was correlated with scores obtained on the AB Word Lists. For each subject, the number of preferences for the two paired-comparison tournaments was pooled and ranked from highest (1) to lowest (7) across the seven hearing aid conditions. These rankings were correlated with the rankings for the AB Word Lists (60 phonemes) using the Spearman rank correlation coefficient.

Table 9 presents the paired-comparison preferences in percent (number of preferences / 36), the phoneme identification test scores in percent, the hearing aid rankings for both measures, and the correlation coefficients for the 8 subjects. The table shows that

Table 9. Paired-Comparison Preferences in Percent (Ranking of Hearing Aids), Phoneme Identification Scores in Percent (Ranking of Hearing Aids), and Spearman Rank Correlation Coefficients.

Paired-Comparison Test (PC1+PC2)

	Subject							
	B	C	F	G	H	I	J	K
HA #								
2	14 (7)	56 (2)	44 (6)	14 (7)	14 (7)	28 (7)	36 (6)	19 (6)
3	44 (5)	67 (1)	47 (4.5)	47 (4)	36 (5)	53 (3.5)	53 (4)	14 (7)
4	36 (6)	44 (1)	47 (4.5)	22 (6)	28 (6)	39 (6)	11 (7)	44 (4.5)
5	58 (4)	53 (3)	42 (7)	28 (5)	56 (4)	50 (5)	44 (5)	44 (4.5)
6	61 (2.5)	50 (4)	50 (2.5)	79 (2)	58 (3)	58 (2)	72 (1)	61 (3)
7	61 (2.5)	33 (7)	50 (2.5)	69 (3)	75 (2)	53 (3.5)	64 (3)	78 (2)
8	75 (1)	47 (5)	69 (1)	92 (1)	83 (1)	69 (1)	69 (2)	89 (1)

Phoneme ID Test (AB Word Lists)

HA #								
2	31 (6)	39 (6)	31 (6)	25 (7)	74 (3)	80 (5)	20 (7)	77 (5)
3	30 (7)	62 (1)	29 (7)	68 (2)	79 (2)	93 (2)	37 (4)	86 (2)
4	41 (5)	54 (3)	32 (5)	44 (6)	42 (6)	58 (7)	28 (6)	54 (6)
5	44 (4)	45 (4)	43 (2)	54 (4)	62 (5)	83 (3)	38 (3)	78 (4)
6	48 (3)	60 (2)	52 (1)	80 (1)	83 (1)	93 (1)	47 (1)	89 (1)
7	59 (1)	38 (7)	33 (4)	47 (5)	35 (7)	59 (6)	31 (5)	53 (7)
8	54 (2)	44 (5)	37 (3)	58 (3)	63 (4)	81 (4)	45 (2)	79 (3)

Spearman Rank Correlation Coefficients

r_s =	.83	.46	.27	.75	-.14	.51	.82	-.10

correlations between paired-comparison judgments and phoneme identification scores were strong (above .80) for Subjects B and J; moderate (between .50 and .80) for Subjects G and I; and weak (below .50) for Subjects C, F, H, and K. The average correlation coefficient for the group of 8 subjects was .49. Because Subjects C and F demonstrated poor test-retest reliability, they were not included in further analyses.

For the 6 subjects who demonstrated good reliability, the mean percent scores on the paired-comparison test and AB Word Lists are displayed in Figure 12. It can be seen that the two sets of scores followed a similar pattern of response across HA#s 2 through 6. The paths intersected at HA#7, however, indicating that preferences increased while phoneme scores decreased for those aids providing the most low-frequency amplification. For all but Subject J, HA#8 ranked highest on the paired-comparison test. For all but Subject B, HA#6 ranked highest on the phoneme identification test.

A more detailed analysis of the phoneme identification scores revealed that the 6 subjects fell into two groupings: Subjects B, G, and J versus Subjects H, I, and K. These data are plotted in Figure 13. It can be seen that the scores averaged for Subjects B, G, and J were lower and less variable than the scores averaged for

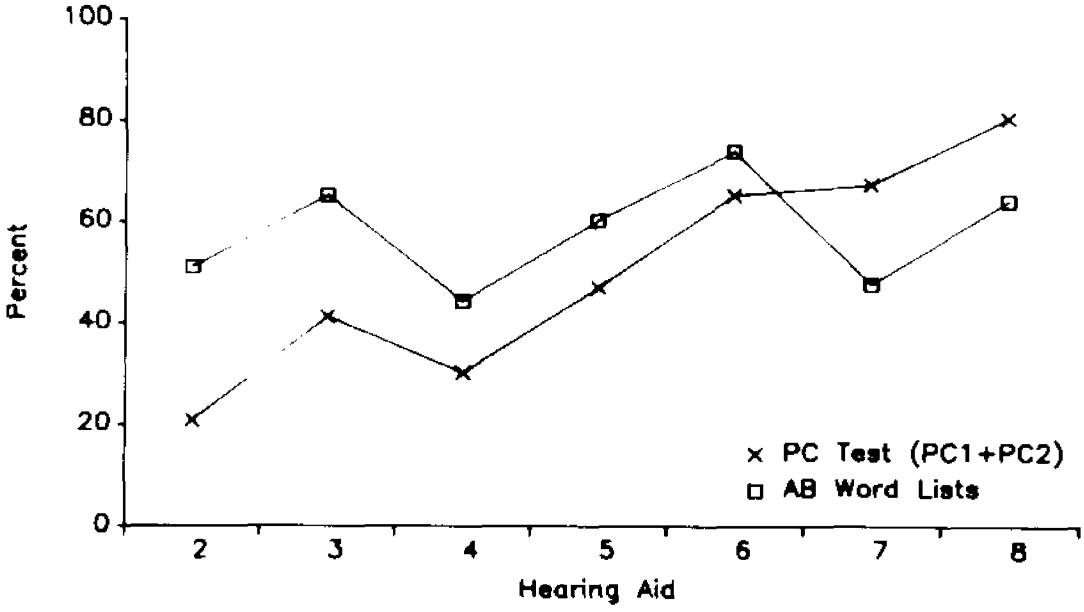


Figure 12. Mean scores obtained by hearing-impaired subjects (n = 6) on the paired-comparison test and AB Word Lists across hearing aids 2 through 8.

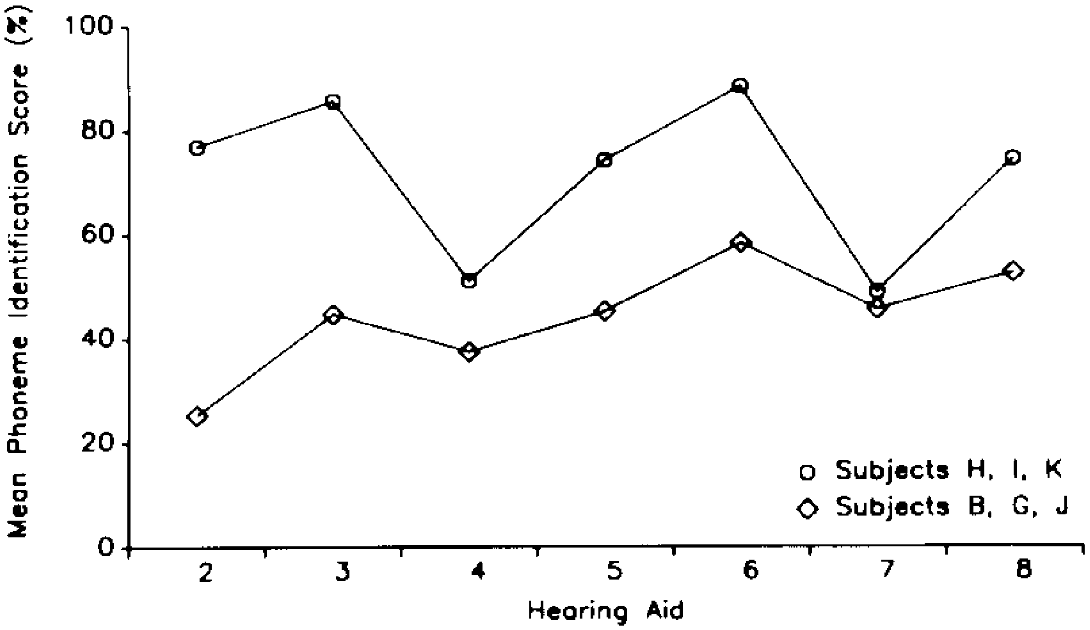


Figure 13. Mean phoneme identification scores on the AB Word Lists for Subjects H, I, and K versus Subjects B, G, and J across hearing aids 2 through 8.

Subjects H, I, and K. As noted earlier, Subjects B, G, and J exhibited moderately strong correlations between paired-comparison judgments and phoneme identification scores, whereas Subjects H, I, and K exhibited weak to moderate correlations.

These two groups also differed by degree of hearing loss and phoneme identification score. With respect to three-frequency average (3FA), Subjects H, I, and K exhibited a mean 3FA of 37 dB HL, whereas Subjects B, G, and J exhibited a mean 3FA of 57 dB HL. This 20-dB difference between groups was found to be statistically significant when analyzed by the Mann-Whitney U Test ($\underline{U} = 0$, $\underline{p} < .05$). With respect to phoneme identification, Subjects H, I, and K, exhibited a mean score of 78.8%, whereas Subjects B, G, and J exhibited a mean score of 71%. This 7.8% difference between groups was not found to be statistically significant.

In a different analysis, individual subject rankings were correlated with group rankings both for paired-comparison judgments and for phoneme identification scores. With respect to the paired-comparison judgments, a close association was observed when individual rankings were correlated with the mean for each of the 6 subjects. Table 10 shows the intercorrelations for paired-comparison

Table 10. Spearman Rank Correlation Coefficients for Paired-Comparison Judgments among 6 Subjects. Preference Judgments were Ranked across Seven Hearing Aids.

Subject	B	G	H	I	J	K
B	--	.96	.99	.92	.87	.86
G		--	.93	.99	.93	.74
H			--	.91	.82	.87
I				--	.92	.66
J					--	.62
K						--

rankings among the 6 subjects. Moderate to strong correlations (above .50) are evident.

The pattern of response for paired-comparison judgments was consistent in that an increase in high-frequency gain resulted in an increase in the number of preferences. More specifically, moving from left to right across rows on the 3 x 3 matrix in Figure 6 (Experiment 2) resulted in higher preference rankings (HA#s 3>2, 5>4, 6>5, 8>7). Further, an increase in low-frequency gain resulted in an increase in the number of preferences. That is, moving from top to bottom down the columns of the matrix resulted in higher preference rankings (HA#s 5>2, 6>3, 7>4, 8>5). Results were ambiguous for those conditions where high-frequency gain was reduced (HA#s 3>4 vs. 6<7).

With respect to phoneme identification scores, results showed that individual subject rankings correlated moderately well with the mean rankings for 5 of the 6 subjects. Subject B exhibited a weak correlation with the mean. The intercorrelations in Table 11 highlight the differences between Subject B and the other 5 subjects. The inconsistencies were isolated to HA#s 3 and 7. Subject B demonstrated low phoneme identification scores for HA#3 and high scores for HA#7. This pattern was

Table 11. Spearman Rank Correlation Coefficients for Phoneme Identification Scores (AB Word Lists) among 6 Subjects. Scores were Ranked across Seven Hearing Aids.

Subject	B	G	H	I	J	K
B	--	.14	-.39	-.11	.36	-.21
G		--	.61	.86	.84	.86
H			--	.82	.42	.89
I				--	.75	.93
J					--	.75
K						--

reversed for the other 5 subjects (see Table 8 for individual scores).

As with the paired-comparison judgments, the response pattern for the group revealed that phoneme identification scores improved with systematic increases in high-frequency gain. In contrast to the paired-comparison judgments, phoneme identification scores did not improve with increases in low-frequency gain. However, scores were seen to diminish when gain was reduced in the high frequencies (HA#s 4<3 and 7<6). This pattern was more obvious for the group composed of Subjects H, I, and K (See Figure 13).

One goal of the previous experiment had been to train hearing-impaired children to make paired-comparison judgments of auditory clarity. The fact that the majority of subjects in Experiment 3 selected a hearing aid that provided low-frequency emphasis suggests that other dimensions of quality may have influenced judgments. To better understand subjects' preferences, performance with personal hearing aids was examined. Phoneme identification scores obtained with personal hearing aids were plotted against scores obtained with aids ranked highest on the combined paired-comparison test. These data are displayed in Figure 14. Subjects placed above

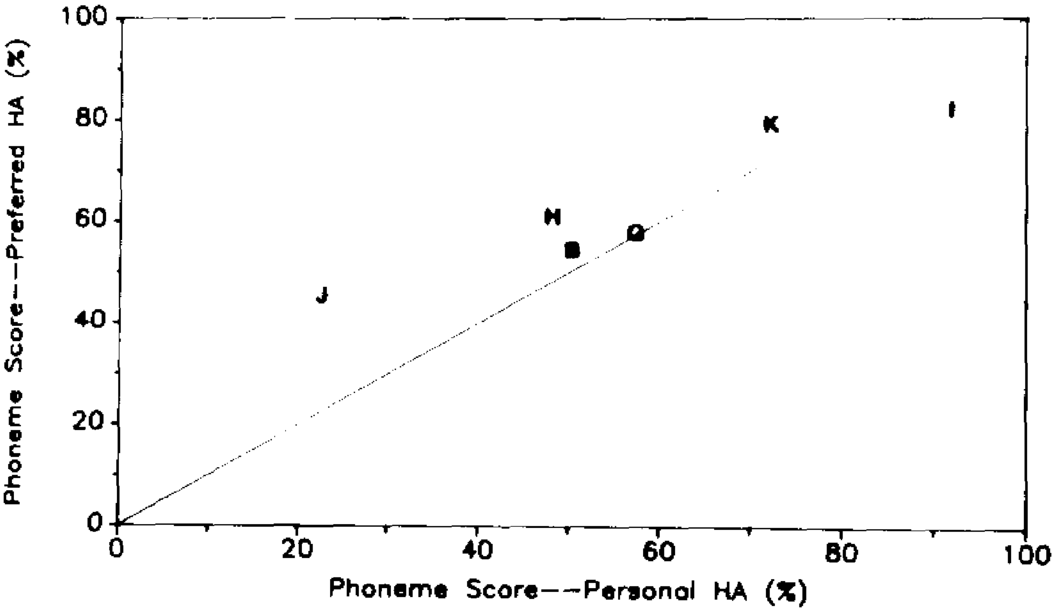


Figure 14. Phoneme identification scores comparing the preferred hearing aid with the personal hearing aid for 6 subjects.

the diagonal performed better with the preferred hearing aid whereas subjects lying below the diagonal performed better with the personal aid. It can be seen that 4 subjects (B, H, J, and K) obtained higher scores with the preferred aid. Of the remaining 2 subjects, Subject G showed no difference between scores, and Subject I obtained a higher score with the personal aid.

Only Subject J demonstrated a statistically significant difference between scores when the data were analyzed on a test of significant differences between proportions approximating a binomial distribution ($\underline{z} = 2.76$, $\underline{p} < .01$). From Figure 14 it can be seen that Subject J exhibited the lowest score with the personal hearing aid. As shown in Appendix B, the personal hearing aid used by Subject J was similar in frequency-gain shape to HA#9, except that the personal hearing aid provided more gain in the mid frequencies. On the paired-comparison test, however, Subject J selected an aid that provided substantially less low- and mid-frequency gain (HA#6).

The other 5 subjects selected HA#8 on the paired-comparison test. As described in Experiment 2, HA#8 increased gain in the low frequencies re: NAL prescription (HA#5), and adjusted gain in the high frequencies to match NAL specifications. Appendix B shows that Subjects B, G, I, and H used personal hearing aids that provided less gain in the high frequencies than would be specified by NAL. In addition, the hearing aid worn by Subject B provided excessive gain in the low and mid frequencies. The personal hearing aid worn by Subject K most closely matched HA#5 in frequency-gain shape. It should be noted that although the subjects were instructed to select the "better" condition of each pair, Subjects H and K both reported that they equated "louder" with "better."

It was also of interest to know whether the hearing aid most preferred by the subjects would degrade speech understanding when compared with a standard prescription. Accordingly, data were plotted comparing phoneme identification scores obtained with the preferred aid against scores obtained with the NAL-prescribed hearing aid (HA#5). These data are presented in Figure 15. It can be seen that 4 subjects (B, G, H, and J) obtained better scores with the preferred hearing aid, 1 subject (K) showed no difference, and 1 subject (I) obtained a better score with HA#5. None of the comparisons was found to reach statistical significance when analyzed on the test of significant differences between proportions.

It should be recognized that the NAL prescription was derived using an input level of 70 dB SPL (Byrne & Dillon, 1986). Because a 60-dB input level was used in the

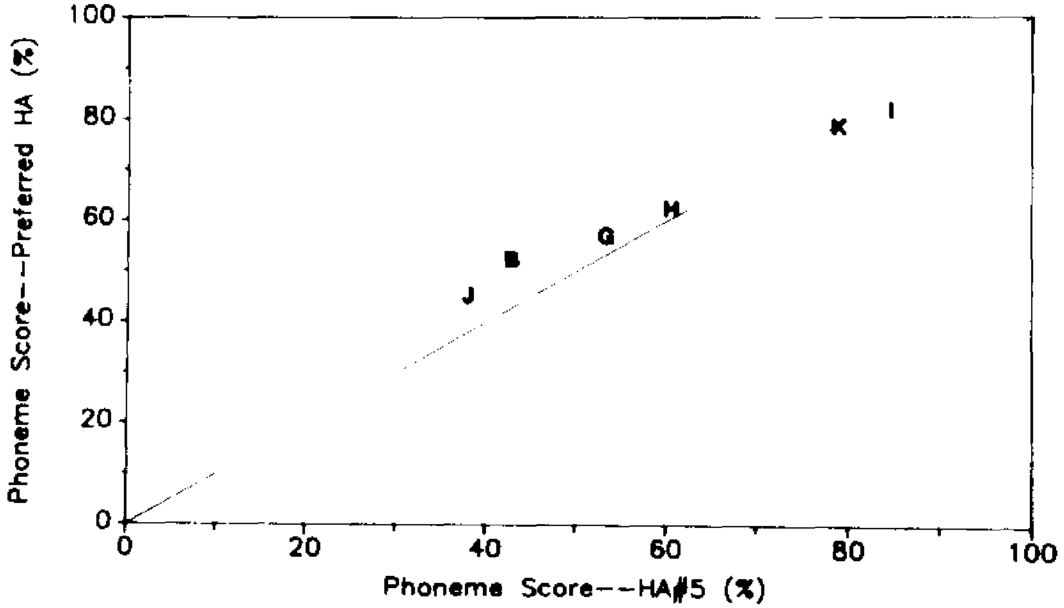


Figure 15. Phoneme identification scores comparing the preferred hearing aid with the NAL-prescribed hearing aid (HA#5) for 6 subjects.

present study, it is probable that performance with HA#5 was poorer than would be expected. Nevertheless, results obtained with the 60-dB input level showed that the preferred aid did not degrade speech intelligibility when compared with HA#5.

Summary

Eight hearing-impaired children judged the clarity of speech processed by seven hearing aids on two paired-comparison, round-robin tournaments. The hearing aids were ranked by number of preferences from highest (1) to lowest (7) for each tournament. The two rankings were correlated using the Spearman rank correlation coefficient. The results showed that correlations between the two sets of rankings were moderate to strong (above .50) for 6 subjects and weak (below .50) for 2 subjects.

Same/different tests were conducted to determine whether inconsistencies in judgments could be attributed to discrimination difficulties. Unfortunately, results were difficult to interpret because some subjects identified the acoustically-same pairs as being different. Another analysis compared the observed variance of phoneme identification scores to the expected binomial variance across hearing aid conditions for individual subjects. Reduced variation among scores was evident for the 2 subjects who exhibited weak correlations between the two paired-comparison tournaments.

The number of preferences for the two paired-comparison tournaments were pooled and ranked from highest (1) to lowest (7) across the seven hearing aid conditions. The rankings were correlated with phoneme identification score rankings. For the 6 subjects who demonstrated good reliability, rank order correlations between paired-comparison judgments and phoneme identification scores were moderately strong (above .70) for 3 subjects, and weak to moderate (below .70) for the other 3 subjects.

Phoneme identification scores obtained with the hearing aid selected by the paired-comparison technique were compared with scores obtained with personal hearing aids and with the NAL-prescribed hearing aid. For 5 of the 6 subjects, scores were either the same or better with the hearing aid selected by the paired-comparison technique.

V. DISCUSSION

Three experiments were conducted to investigate the feasibility of using a paired-comparison technique to select hearing aids for children. The purpose of Experiment 1 was to determine the age at which at least 50% of normal-hearing children could be expected to meet criterion on a paired-comparison task involving judgments of auditory clarity. A training/testing protocol presented pairs of conditions that contrasted visual, auditory-visual, and auditory clarity in a videotaped children's story.

Twenty-five normal-hearing children between the ages of 4.0 and 6.4 years participated in Experiment 1. Five age groups were formed with 5 subjects per group. The youngest age group in which all 5 subjects passed the auditory paired-comparison task determined the age at which no less than 50% of children in the population could be expected to perform the task at the .05 probability level. That is, if 50% of children in the population can pass the task, then the probability of 5 out of 5 "passers" occurring by chance is less than .05.

It is recognized that use of a small sample size increases the risk of committing a Type II error. For example, if 90% of children in a population age group can pass the task, then the probability of 5 out of 5 subjects in the study passing is .59. The error probability is high (.41) that less than 5 subjects will pass. On the other hand, the probability of making a Type I error is reduced because all 5 subjects in the sample are required to pass the task.

Results from Experiment 1 showed that performance was related to age. With each increment in age group, a greater number of children passed the different paired-comparison tasks. Further, the proportion of correct responses across test pairs increased with increments in age group when the data were collapsed across the three paired-comparison tasks. The youngest age group in which all 5 children passed the auditory task was 6.0 to 6.4 years.

The age at which the children passed the different paired-comparison tasks was in accordance with the developmental stages documented in the cognitive and linguistic literature which assert that the ability to make comparative judgments matures between the ages of 4 and 7 years (Gitterman & Johnston, 1983; Inhelder & Piaget, 1964; Piaget, 1967). In the present study, children as young as 4.5 years were able to pass some or all of the paired-comparison tasks. This finding provides evidence to support the notion that performance closely estimated ability.

The children in the 4.0-year age group were not able to be trained on the paired-comparison task. However, data were presented to suggest that at least 1 child in this age group was able to perform the paired-comparison task, but was unable to make the finer discriminations required to progress from training to testing.

On the other hand, all children in the 4.0-year age group were able to perform the simpler task of selecting the clearer condition from among matched pairs of clear and blurry photographs. Optometrists have also reported that children 5 years of age (and occasionally younger) can perform this type of task.

It has been reported by Miller (1979) that children are able to extract the distinctive feature when forced to make a two-choice decision regarding dimensional attributes. In the present study, the youngest children were able to extract the feature of clarity from the pair of matching photographs. However, it is not known whether they would have achieved the same success had the photograph test presented pairs that required finer discriminations.

It is also unknown whether the younger children would have achieved greater success on the training/testing protocol had the stimulus differed only in clarity, as it did with the clear and blurry photographs. The continuous discourse stimuli used in the actual experiment varied in content as well as in clarity. Consequently, the younger children may have experienced greater difficulty in extracting the feature of clarity amidst the other variations occurring at the same time. It is possible that use of paired sentences (or words) rather than continuous discourse might have simplified the task for these children.

Results from Experiment 1 also showed that performance on the auditory task paralleled performance on the visual task. With the exception of 2 children, the 18 children who passed the visual task also passed one or both of the two auditory tasks. In addition, the proportion of correct responses across test pairs was fairly well-matched for the visual, auditory-visual, and auditory tasks when the data were collapsed across subjects. Such findings imply that the paired-comparison task was not modality dependent. However, the relative contribution of the visual component in training children to perform the auditory task is unknown.

It is important to bear in mind that the number of children in this experiment was small, specifically 5 children per group. However, as noted earlier, even with a small sample size the probability of all 5 children in age group reaching criterion is statistically significant (<.05) if as many as one half of children at that age in the general population are at criterion level. Although the small sample size does not allow for a precise estimate of the proportion of children in each age group that can be expected to reach criterion level, it does indicate that a substantial proportion (probably the majority) of children in the population are likely to reach it when all 5 children in a corresponding age group reach criterion. It should be noted that the power of the statistical test can be increased by pooling data of adjacent age groups (e.g., by forming groups of 10 children spaced a year apart). This technique, however, reduces the precision with which the expected age can be specified for a majority of the children to reach criterion level.

In summary, Experiment 1 showed that the youngest age group in which all 5 normal-hearing children met criterion on the auditory paired-comparison task was between 6.0 and 6.4 years. Although children as young as 4.5 years of age were able to pass some or all of the tasks in the training/testing protocol, it is possible that the younger children might have passed more tasks had the activity been easier to perform. The paired-comparison task was not found to be modality dependent.

The purpose of Experiment 2 was to determine the age at which at least 50% of hearing-impaired children could be expected to meet criterion on a paired-comparison task involving judgments of auditory clarity. Although hearing-impaired children do not represent a homogeneous population, the children in the present study were selected on the basis of communication mode (oral), hearing aid experience (6 months or more), and phoneme identification score (50% or better).

Ten English-speaking children with mild to moderately-severe sensorineural hearing loss participated in Experiment 2. Their ages ranged between 5.5 and 7.4 years. Two age groups were formed with 5 subjects per group; 5.5 to 6.4 years, and 6.5 to 7.4 years.

The results showed that all 10 children passed the visual task and one or both of the two auditory-visual tasks. However, while all 5 children in the older age group passed one or both of the two auditory paired-comparison tasks, only 2 children in the younger group passed this task. Further, the older age group obtained a higher proportion of correct responses across test pairs when the data were collapsed across the three paired-comparison tasks.

In comparing performance between the normal-hearing and hearing-impaired children, two observations may be noted. First, there was no apparent difference between the two sample groups with respect to performance on the visual paired-comparison task. All children 5.5 years and older were able to pass this task. Such findings support assertions put forth by Furth (1964) that normal-hearing and hearing-impaired children are able to achieve similar levels of performance on some cognitive tasks, particularly those tasks that do not require language.

With respect to language, it is of interest to note that results on the receptive language test (particularly the inflectional morphemes) did not reflect performance on the paired-comparison task. That is, all of the children demonstrated comprehension of the inflectional morphemes. However, it is also noteworthy that both groups of subjects performed at or above age level on the TACL.

Findings from Experiment 2 support the notion that hearing-impaired children are cognitively able to perform a paired-comparison task that involves judgments of visual clarity. This ability appears to be independent of the hearing impairment, although the influence of receptive language age is unknown.

A second observation is that the hearing-impaired children experienced greater difficulty performing the auditory paired-comparison task than did the normal-hearing children. Auditory low-pass filtering was used with the normal-hearing children whereas frequency-gain shaping was used with the hearing-impaired children. Results indicated that the frequency-gain shapes were not well-matched with the visual conditions. A lower proportion of correct responses across test pairs was obtained for the auditory conditions relative to the visual conditions when the data were collapsed across subjects.

Because of the mismatch between the visual and auditory conditions, it is not clear whether the auditory paired-comparison task was cognitively more difficult for the hearing-impaired children, or whether performance was influenced primarily by the hearing impairment. Post hoc analyses on phoneme identification scores provided evidence to suggest that some of the auditory conditions were not significantly different for the children as a group. This finding implies that discrimination difficulties contributed to these results.

To summarize, Experiment 2 showed that all 5 hearing-impaired children between the ages of 6.5 and 7.4 years passed the auditory paired-comparison task, whereas only 2 of the 5 children between 5.5 and 6.4 years of age passed the auditory task. In contrast, all 10 children passed the visual and auditory-visual tasks. The relatively poorer performance exhibited on the auditory task was attributed in part to the effects of hearing impairment.

The purpose of Experiment 3 was to investigate whether a paired-comparison technique could be used effectively to select appropriate hearing aids for hearing-impaired children who were able to perform the auditory paired-comparison task. Eight hearing-impaired children participated. Their ages ranged from 5.7 to 7.8 years, with a mean age of 6.7 years.

The children were evaluated on two paired-comparison, round-robin tournaments to judge auditory clarity of seven hearing aid, frequency-gain shapes. The number of preferences per hearing aid condition were ranked for each test, and the two sets of rankings were correlated for each subject. The correlations were used as a measure of test-retest reliability.

Analysis of test-retest reliability showed that correlations between the two sets of hearing aid rankings were moderate to strong (above .50) for 6 children and weak (below .50) for 2 children. Using this measure, the data were considered to be reliable for the 6 children who demonstrated correlations above .50.

These results are consistent with findings reported for hearing-impaired adults; paired-comparison judgments were shown to be reliable even though some variability existed. With adults, performance variability was attributed to an inability to perceive differences among conditions and/or to changes in judgmental criterion (Levitt, Sullivan, Neuman, & Rubin-Spitz, 1987; Punch & Parker, 1981).

To determine whether children's inconsistencies could be explained by discrimination difficulties, same/different tests were conducted. Unfortunately, results on these tests were difficult to interpret because some subjects identified the acoustically-same pairs as being different (again, possibly due to the varying nature of continuous discourse). Alternatively, a different analysis was performed comparing the observed variance of phoneme identification scores to the expected binomial variance across hearing aids for each subject. The comparisons were not found to be significantly different for the 2 children who demonstrated poor reliability. This finding suggested that an inability to perceive differences among conditions may have contributed to the variability. Performance variability was also attributed to limitations in attention span for these 2 children.

In the study conducted by Levitt, Sullivan, Neuman, and Rubin-Spitz (1987), a 6-dB/octave slope was used to vary the frequency and gain of nine hearing aid conditions. For most of their subjects, the observed standard deviation of speech discrimination scores did not exceed the expected binomial standard deviation across the nine conditions. Levitt et al. (1987) concluded that the range of variation among conditions was poor relative to the test-retest variability of the scores. Because the adult subjects were allowed to adjust the output level for each condition, the investigators speculated that this adjustment may have reduced the perceptual differences among conditions.

The present study differed from the Levitt, Sullivan, Neuman, and Rubin-Spitz (1987) study in that the frequency-gain shapes were varied in 12-dB steps and were presented at a fixed input level of 60 dB L_{eq}. The observed variance was found to be significantly greater than the expected variance for 5 of the 6 children who demonstrated good reliability. This finding suggested that the range of variation among conditions was such that most of the children were able to perceive differences among conditions.

To measure the relationship between paired-comparison judgments and speech intelligibility, results on the paired-comparison test were correlated with phoneme identification scores. The number of preferences for the two paired-comparison tournaments was pooled and ranked across the seven hearing aids. These rankings were correlated with phoneme identification score rankings for each subject.

Results showed that the 2 children who demonstrated poor test-retest reliability, also demonstrated weak correlations (below .50) between paired-comparison judgments and phoneme identification scores. For the 6 children who demonstrated good reliability, the correlations between preference judgments and phoneme identification scores were found to be moderately strong (above .70) for 3 children, and weak to moderate (below .70) for the other 3 children. The children who demonstrated the stronger correlations exhibited poorer hearing thresholds and phoneme identification scores.

The hearing aid selected by most of the hearing-impaired children was one that provided low-frequency emphasis. Despite the risk of upward spread of masking, studies have demonstrated that low-frequency gain does not necessarily degrade speech understanding (e.g., Punch & Beck, 1986). A similar conclusion was reached in the present study. Phoneme identification scores obtained with the preferred hearing aid did not differ significantly from scores obtained with the NAL-prescribed hearing aid.

In a study that explored adults' ratings of perceived sound quality, Gabrielsson et al. (1988) found that clarity was associated with an increased high-frequency response and either a flat or decreased low-frequency response for hearing-impaired listeners. In the present study, the hearing aid selected by the paired-comparison technique delivered more high-frequency gain than was typically provided by the children's personal hearing aids. This finding implies that judgments may have been based on clarity.

On the other hand, because the selected hearing aid also provided low-frequency emphasis for most of the children, loudness may have provided the more salient cue on which judgments were based. In the Gabrielsson et al. (1988) study, it was speculated that loudness may be the more familiar attribute for hearing-impaired listeners. In the present study, the 2 children who exhibited weak correlations between preference judgments and phoneme identification scores admitted that they selected the louder condition within each pair. These 2 children were from the group of 3 children who exhibited better thresholds and phoneme identification scores. Because hearing was quite good for these children, it is possible that the hearing aid conditions provided cues that were more salient in loudness than in clarity.

The 3 children showing the stronger correlations between paired-comparison judgments and phoneme identification scores were from the group that exhibited poorer thresholds and phoneme identification scores. Because of poorer hearing, some conditions may have been very difficult for these children to perceive and/or understand. As such, increases in loudness may have contributed to improved clarity. It is possible that judgments were based both on loudness and clarity for these children.

In summary, Experiment 3 showed that correlations between two paired-comparison tournaments were moderate to strong (above .50) for 6 of 8 hearing-impaired children. For the 6 children who demonstrated good reliability, correlations between paired-comparison judgments and phoneme identification scores were moderately strong (above .70) for the 3 children who exhibited the poorer hearing, and weak to moderate (below .70) for the 3 children who exhibited relatively better hearing. It is speculated that judgments were based on clarity and/or loudness. The hearing aid selected by the paired-comparison technique did not degrade speech intelligibility when compared with the NAL-prescribed hearing aid. It is concluded that a paired-comparison technique can be used reliably to select hearing aids for some hearing-impaired children.

Results from this study are of clinical relevance. In fitting hearing aids on adults, audiologists typically start with a standard prescription and readjust the setting according to patient preference. With children, hearing aids are readjusted only when discomfort is shown or sound not detected. This study confirms that hearing-impaired children 6.5 years of age and occasionally younger are able to make reliable judgments that relate in some way to speech intelligibility (as evidenced by the high correlations existing between paired-comparison judgments and phoneme identification scores for some children). Such children should be able to play a more active role in the selection process.

Results from this study also imply that children may perform better with the hearing aid selected by the paired-comparison technique than with an aid selected by a standard prescription. This finding suggests that the paired-comparison technique can better individualize the fitting over the more traditional approach and, if used, could lead to better acceptability of the aid by the child. Should a discrepancy occur between the aid selected by the paired-comparison technique and that recommended by the prescription, the clinician may be faced with a dilemma. Results from this study and from the literature suggest that the aid selected by the paired-comparison technique would be the preferable choice because of better face validity. However, should the aid selected by the paired-comparison technique seem inappropriate for the child given the characteristics of the hearing loss, the clinician would want to explore the possible causes contributing to the child's selection.

Certain cognitive and linguistic abilities as well as listening skills and auditory capacity contribute to the hearing-impaired child's ability to perform auditory paired comparisons. The extent to which these factors contribute can be estimated from the control provided by the normal-hearing group and the visual paired-comparison task. The normal-hearing children performed at age levels consistent with the cognitive and linguistic literature for tasks involving comparative judgments. Results with normal-hearing children provide a reference for comparing performance with hearing-impaired children on similar tasks. The second control, the visual paired-comparison task, provided evidence to suggest that hearing-impaired children of specific chronologic and language ages were cognitively able to perform a paired-comparison task.

The last two factors, listening skills and auditory capacity, affected performance by influencing children's ability to perceive differences among hearing aids. These are the factors of interest in fitting a hearing aid to a child. Once a hearing aid has been prescribed, it is possible to hone listening skills with auditory training such that an improved fitting of the hearing aid can be achieved at a later stage (i.e., an initial fitting either by prescription or by the paired—comparison technique can be used to obtain an approximate fit which can be improved using the paired-comparison technique following a period of training).

Standard hearing aid selection procedures presently used with children include the prescriptive approach (e.g., Seewald & Ross, 1988) and the comparative approach (e.g., Matkin, 1986; Northern & Downs, 1984; Schwartz & Larson, 1977b). These approaches rely on limited information to provide the data needed to select or fit a hearing aid. Such information includes behavioral observation, threshold tests, functional or insertion gain, and discrimination score. Guidelines are provided that enable the clinician to amplify speech to audible levels and to estimate the maximum power output.

On the other hand, standard techniques in use today are inadequate for adjusting the many parameters available with new digital hearing aids and other sensory devices such as cochlear implants, tactile aids, and assistive listening devices. As sensory aids continue to improve and provide a wider range of options to the clinician, improved methods of selection will be needed so that children can benefit from these technological advances. The paired-comparison technique is applicable for evaluating a variety of parameters. As such, it offers the advantage of flexibility over other existing methods of selection.

Recent studies have demonstrated that children are capable of performing certain auditory tasks that were once believed to be too difficult. For example, Kawell et al. (1988) and Macpherson et al. (1989) demonstrated that it is possible to obtain reliable loudness discomfort levels in children as young as 5 years of age when techniques are implemented that account for children's linguistic and cognitive abilities.

The present study has similarly shown that a paired-comparison technique can be adapted for use with children. Special procedures were implemented to accomplish this goal. First, the word "better" was substituted for the word "clearer" in instructing the child to perform the task. Secondly, the task was made more interesting by use of a videotaped children's story. Thirdly, various activities were available to ensure that the child understood how to perform the paired-comparison task.

Finally, the tester was seated next to the child throughout training and testing to provide instruction as needed.

Although these techniques were shown to be helpful, additional modifications are still warranted. For example, not all children were successful in performing the paired-comparison task. Before the paired-comparison technique can be implemented clinically, it is important to determine which children are capable of performing the task. As noted earlier, it is likely that the task could be administered to normal-hearing children as young as 6 years of age or even younger. However, one should not rely on the average statistical characteristics of the population in order to decide on whether or not to use the procedure (i.e., it is conceivable, for example, that a few children as young as 4 years of age may be able to perform the paired-comparison task and a few children older than 6 years of age may not be able to perform the paired-comparison task). It is important, therefore, to have a reliable screening instrument to ensure that the child can perform the task.

The visual paired-comparison task appeared to be a good screening tool. The data showed that almost all children who passed the visual paired-comparison task also passed the auditory paired-comparison task (i.e., of the

18 normal-hearing children who passed the visual paired-comparison task, 16 of these children passed the auditory paired-comparison task). All of the children who were unable to pass the visual task did not pass the auditory task. These results suggest that the visual paired-comparison task can be used as a screening instrument provided it is refined such that all who can pass the visual paired-comparison task can be expected to perform the auditory paired-comparison task.

For those children unable to pass the screening, special procedures might be implemented to simplify the task. As suggested earlier, use of words or sentences instead of continuous discourse might better facilitate understanding of the paired-comparison task.

For the children who demonstrated weak correlations between judgments of clarity and phoneme identification scores, an increase in training and/or experience with the concept of clarity might be advantageous. Trabasso and colleagues showed that training was important in determining young children's ability to make transitive inferences (Bryant & Trabasso, 1971; Riley & Trabasso, 1974). In those studies, children were trained in three or four short sessions to discriminate among pairs of sticks that differed in length and color. The training involved associating length with color, presenting pairs in reversible order, requiring the children to indicate which was "taller" or "shorter," and providing feedback.

Similar techniques were used in the present study. Auditory clarity was associated with visual clarity, and feedback was provided during the training component of the protocol. However, in the present study, the word "better" was substituted for the word the "clearer." Had the terms "clearer" or "blurrier" been applied in a consistent manner, the correlations between preference judgments and phoneme identification scores might have been stronger. On the other hand, the amount of training described in the Trabasso studies would not be clinically practical.

An alternate solution might be to provide hearing aid experience prior to testing, as demonstrated by Byrne and Cotton (1988). They showed that a high proportion of hearing-impaired adults selected the new NAL prescription on a paired-comparison test following 2 to 3 weeks of experience with that fitting. The NAL prescription served as the reference condition from which other variations were compared in the paired-comparison test. However, even after the trial period, some subjects still selected a variation in prescription. These findings underscore the importance of providing experience with the reference condition prior to performing a test that requires listener judgments. Experience could help delineate children's preferences with respect to attributes of sound quality such as loudness, clarity, and so forth.

Although results from this study are promising, further research is needed. For example, it is unknown whether the training/testing protocol used in the present study was essential to successful performance on the paired-comparison task. Additional studies would be needed to determine whether the methods used in training were mandatory, or whether alternate methods would be superior.

The hearing aids evaluated in the present study varied by frequency-gain shape. This parameter was chosen in an attempt to replicate studies conducted with adults. However, other characteristics or stimuli should be investigated. Indeed, it would be interesting to know whether children would have selected a different hearing aid had noise been introduced during testing. Moreover, other dimensions of perceived sound quality would be most interesting to investigate both with normal-hearing and hearing-impaired children. With these and other studies, the paired-comparison technique could someday enable researchers and clinicians to test children on a number of hearing aid conditions.

In conclusion, the goal of this study was to determine the feasibility of using a paired-comparison technique to select hearing aids for children. Results indicated that a paired-comparison technique is feasible for use with hearing-impaired children 6.5 years of age (and occasionally younger) who are capable of performing the task reliably. These results support use of the paired-comparison technique to select hearing aids for children, and justify future studies to identify other more effective methods of implementing the technique.

APPENDIX A

The paired-comparison photograph test was constructed from the DLM sequential picture cards (DLM Teaching Resources, Allen, Texas). Pictures from the six-card series were reproduced in clear and blurry color photographs. A clear and blurry example from the test is shown below.

APPENDIX B

Appendix B presents hearing aid output levels with corresponding phoneme identification scores for individual subjects. Output levels for puretones were measured at an input level of 60 dB SPL (re: 20 µPa). The nine frequency-gain filter shapes were measured at the headphone through a 6-cm^3 coupler using a Bruel & Kjaer sound level meter (type 2231). Personal hearing aids were measured through a 2-cm^3 coupler on the Fonix Hearing Aid Test System (Frye Electronics, Inc.).

Levels for personal hearing aids were corrected for earmold insertion loss using the Shaw (1974) and Shaw and Vaillancourt (1985) freefield to eardrum transformation curves. It is recognized that differences in ear canal resonance exist between children and adults. However, for children of the ages tested in the present study, coupler to real ear differences have not been shown to differ significantly when compared to similar measures obtained on adults (Barlow, Auslander, Rines, & Stelmachowicz, 1988).

APPENDIX B

Output levels for nine hearing aid filter shapes (HA) and personal hearing aids (PA)[a] in dB SPL re: 20 µPa, and phoneme identification scores on the AB Word Lists (ABW).

Subject A

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	62.4	71.2	82.1	91.0	84.2	78.0	77.1	74.5	10.83
2	62.6	71.2	82.2	91.3	91.5	90.1	89.1	86.6	17.50
3	62.6	71.0	82.1	91.3	98.8	101.8	100.8	98.4	18.33
4	74.5	83.1	86.8	91.3	84.1	77.9	77.0	74.5	7.50
5	74.5	83.1	86.9	91.5	91.5	90.1	89.0	86.5	19.17
6	74.4	82.9	86.8	91.5	98.7	101.9	100.8	98.4	25.00
7	86.5	94.8	91.6	91.6	84.1	77.9	77.0	74.4	24.17
8	86.4	94.7	91.6	91.6	91.3	89.9	88.8	86.4	23.33
9	86.4	94.7	91.7	91.6	98.7	101.9	100.8	98.4	40.00
PA	75.5	86.2	91.4	97.4	95.9	91.5	88.1	82.7	53.33

---

[a] For the personal hearing aid:
.25 kHz = .2 kHz
3.15 kHz = 3.2 kHz

Subject B

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	47.9	53.2	73.3	83.3	78.8	74.8	75.4	79.2	38.33
2	47.7	52.9	73.2	83.4	86.0	87.0	87.3	91.1	30.83
3	46.2	52.1	72.8	83.1	93.1	99.0	98.8	103.2	30.00
4	59.6	64.8	78.0	83.6	78.7	74.8	75.4	79.2	40.83
5	59.6	64.6	78.0	83.7	86.0	87.0	87.2	91.0	43.33
6	59.4	64.4	77.9	83.7	93.2	98.8	99.0	102.8	48.33
7	71.7	76.7	82.8	83.9	78.9	74.8	75.3	79.1	59.17
8	71.6	76.5	82.7	83.8	85.9	86.9	87.2	91.0	54.17
9	71.5	76.4	82.7	83.8	93.2	98.7	99.1	102.9	52.50
PA	79.5	86.7	87.4	89.9	95.9	97.0	79.6	81.7	50.00

Subject C

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	47.6	56.1	73.9	84.5	72.7	57.6	58.9	57.7	37.50
2	47.7	56.1	73.7	84.9	80.1	69.8	70.9	69.8	39.17
3	47.7	56.1	73.8	85.2	87.4	82.1	82.9	81.8	61.67
4	59.6	68.2	78.5	85.1	72.8	57.6	58.9	57.7	54.17
5	59.6	68.0	78.5	85.3	80.1	69.8	70.9	69.8	45.00
6	59.6	68.1	78.5	85.5	87.4	82.1	82.8	81.8	60.00
7	71.7	79.9	83.3	85.3	72.8	57.6	58.9	57.7	38.33
8	71.7	79.9	83.3	85.6	80.1	69.8	70.9	69.8	44.17
9	71.8	79.8	83.3	85.7	87.3	82.1	82.8	81.8	60.00
PA	62.5	72.7	77.4	82.4	75.4	68.0	64.6	61.7	63.33

Subject D

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	48.5	60.2	76.2	85.8	75.1	65.3	66.1	65.1	17.50
2	48.5	60.2	76.1	86.1	82.7	77.4	78.1	77.1	53.33
3	48.6	60.4	76.1	86.1	89.8	89.5	90.0	89.1	82.50
4	60.5	72.0	80.8	86.3	75.1	65.2	66.0	65.0	63.33
5	60.4	72.0	80.8	86.2	84.0	77.4	78.0	77.0	70.00
6	60.3	72.0	80.8	86.6	89.8	89.4	89.9	89.0	88.33
7	72.4	83.8	85.5	86.6	78.1	65.1	66.0	64.9	50.83
8	72.4	83.8	85.6	86.7	86.2	77.3	78.0	77.0	85.83
9	72.3	83.8	85.6	86.3	89.6	89.4	89.9	89.0	76.67
PA	73.0	82.7	88.4	95.4	91.4	89.0	86.1	85.2	74.17

Subject E

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	47.5	55.2	72.3	81.1	76.1	71.1	68.3	64.7	19.17
2	47.5	55.4	72.3	81.4	83.4	83.2	80.3	76.8	51.67
3	47.3	55.2	72.3	81.5	90.7	95.0	92.1	88.6	57.50
4	59.3	67.1	77.0	81.5	76.1	71.1	68.3	64.7	34.17
5	59.0	67.1	77.1	81.7	83.4	83.2	80.3	76.7	53.33
6	59.4	67.2	77.2	81.8	90.8	95.2	92.4	88.7	77.50
7	71.3	78.4	81.7	81.7	75.9	70.9	68.2	64.6	65.83
8	71.4	79.0	81.9	81.9	83.4	83.3	80.3	76.7	58.33
9	71.4	79.0	82.0	82.0	90.8	95.2	92.3	88.9	63.33
PA	57.0	65.2	72.9	80.4	79.9	77.0	75.1	75.2	47.67

Subject F

	Frequency (kHz)								Phoneme ID
HA#	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
1	52.8	65.7	81.4	90.1	83.3	77.1	77.9	76.8	15.00
2	52.8	65.8	81.4	90.5	90.9	89.3	89.9	88.9	30.83
3	52.3	65.7	81.4	90.7	98.1	101.4	101.6	100.6	29.17
4	64.4	77.5	86.0	90.5	84.4	77.2	77.8	76.6	31.67
5	64.4	77.6	86.1	90.7	91.2	89.3	89.8	88.8	43.33
6	64.6	77.5	86.1	90.5	97.9	101.0	101.5	100.7	51.67
7	76.4	89.3	90.8	90.5	83.1	77.1	77.7	76.7	33.33
8	76.4	89.3	90.9	90.8	90.5	89.2	89.7	88.7	36.67
9	76.3	89.2	90.8	90.8	97.8	101.0	101.5	101.5	46.67
PA	65.5	81.7	92.4	100.4	90.9	89.0	79.1	80.7	42.50

Subject G

	Frequency (kHz)								Phoneme ID
HA#	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
1	58.8	69.2	82.9	90.4	82.5	77.3	74.8	72.6	27.50
2	58.9	69.2	82.9	90.8	88.5	89.6	86.8	84.3	25.00
3	58.9	69.3	83.0	91.0	95.9	101.3	98.7	96.6	67.50
4	71.0	81.1	87.6	90.9	81.1	77.5	74.8	72.5	44.17
5	71.0	81.1	87.7	91.1	88.4	89.7	86.8	84.6	54.17
6	71.0	81.4	87.9	91.2	95.9	101.5	98.8	96.6	80.00
7	83.1	93.0	92.6	91.2	81.2	77.5	74.9	72.6	46.67
8	83.1	93.0	92.6	91.3	88.5	89.6	86.8	84.6	58.33
9	83.2	93.1	92.7	91.5	95.9	101.5	98.8	96.6	75.83
PA	65.0	76.7	81.4	86.9	84.9	80.0	75.6	77.7	57.50

Subject H

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	47.2	50.7	63.0	72.1	74.1	72.4	70.7	71.5	45.00
2	47.0	50.7	63.0	72.3	81.5	84.7	82.8	83.0	74.17
3	46.5	51.1	63.1	72.5	88.8	96.4	94.8	94.9	79.17
4	59.1	62.6	67.7	72.4	74.0	72.4	70.7	70.9	41.67
5	59.1	62.5	67.8	72.5	81.4	84.7	82.8	83.0	61.67
6	58.9	62.4	67.9	72.5	88.8	96.6	94.8	95.0	83.33
7	71.1	74.4	72.6	72.5	74.0	72.7	70.7	70.9	35.00
8	71.1	74.3	72.6	72.7	81.4	84.7	82.8	82.9	63.33
9	71.1	74.3	72.6	72.6	88.8	96.6	94.7	94.9	75.00
PA	53.0	55.7	62.9	72.4	66.4	62.5	72.1	58.2	48.33

Subject I

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	48.4	57.3	70.3	78.9	68.8	62.8	60.4	58.0	35.00
2	48.4	56.5	70.3	79.2	75.2	74.9	72.3	70.0	80.00
3	48.3	57.2	70.3	79.4	82.5	86.9	84.3	82.0	92.50
4	60.3	68.2	75.0	79.3	68.6	62.8	60.3	58.0	57.50
5	60.3	68.2	74.9	79.5	75.1	74.8	72.3	70.0	83.33
6	60.3	68.1	75.0	79.6	82.5	86.9	84.3	82.0	93.33
7	72.2	79.9	79.7	79.6	67.8	62.8	60.3	58.0	59.17
8	72.2	79.9	79.8	79.7	75.1	74.8	72.3	69.9	80.83
9	72.2	80.0	79.9	79.5	82.5	81.9	84.3	82.0	97.50
PA	55.5	64.3	72.9	81.9	78.4	71.5	61.6	56.7	88.33

Subject J

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	55.6	65.6	77.7	85.4	78.1	70.9	70.4	72.0	12.50
2	55.6	65.6	77.8	85.7	85.5	83.0	82.5	84.1	20.00
3	55.6	65.7	77.8	85.9	94.9	94.8	94.4	96.0	36.67
4	67.4	77.5	82.5	85.3	78.1	70.9	70.4	72.0	27.50
5	67.4	77.5	82.6	86.0	85.5	83.1	82.4	84.1	38.33
6	67.3	77.4	82.6	86.0	92.8	94.9	94.3	96.0	46.67
7	79.5	89.3	87.3	86.1	78.2	79.0	70.5	72.0	30.83
8	79.5	89.3	87.4	86.1	85.5	83.1	82.4	84.1	45.00
9	79.5	89.3	87.4	86.1	92.8	94.9	94.3	96.0	59.17
PA	78.0	88.2	96.4	105.4	96.4	92.5	90.6	80.2	22.50

Subject K

	Frequency (kHz)								Phoneme ID
	.25	.5	.8	1	1.6	2	3.15	4	ABW (%)
HA#									
1	47.8	53.0	67.1	77.2	72.1	66.4	65.2	63.9	59.17
2	48.1	53.0	67.1	77.5	79.6	78.9	77.2	76.0	76.67
3	48.0	52.9	67.1	77.7	86.9	91.1	89.2	87.9	85.83
4	59.9	64.9	71.8	77.5	72.1	66.5	65.1	63.9	54.17
5	59.9	65.1	71.9	77.8	79.5	78.9	77.2	75.9	78.33
6	59.8	64.9	71.9	72.9	86.9	91.2	89.1	87.9	89.17
7	71.8	76.7	76.6	77.8	72.1	66.4	65.1	63.9	52.50
8	71.9	76.8	76.7	78.0	79.5	78.9	77.2	75.9	79.17
9	71.8	76.7	76.7	78.1	86.9	91.1	89.1	87.9	86.67
PA	59.5	66.2	75.4	84.4	75.4	77.0	71.1	67.7	71.67

APPENDIX C

Loudness Discomfort Level (LDL) Procedure

(Modified from Macpherson et al., 1989)

1) The child observes the tester hit a button that causes a red light to come on. After the light is turned on and off, the tester begins to move a toy car along the table. The red light is turned on again and the tester stops the car while saying, "red light--time to stop." The tester and child take turns moving the toy car and operating the red light.

2) The tester fills a container with checkers. When the container is filled, the tester hits the red light button and says, "no more--time to stop." The tester and child take turns filling the container and operating the red light.

3) The child calls out the tester's name. The tester asks the child to raise his or her voice. When the voice becomes too loud, the tester points to her ear and then hits the red light button while saying, "no more--time to stop."

4) The child listens under headphones to a lady reading a story. The volume is gradually increased. The child is instructed to hit the red light button and say "no more--time to stop" when the speech becomes too loud.

BIBLIOGRAPHY

American National Standards Institute (1989). American National Standard specification for audiometers (ANSI S3.6-1969). New York: ANSI.

American Speech-Language-Hearing Association (1985). Guidelines for identification audiometry. Asha, 27, 49-52.

American Speech-Language-Hearing Association (1990). Guidelines for screening for hearing impairment and middle-ear disorders. Asha Supplement, 32, 17-24.

Anderson, D. R., Alwitt, L. F., Lorch, E. P., & Levin, S. R. (1979). Watching children watch television. In G. A. Hale & M. Lewis (Eds.), Attention and cognitive development (pp. 331-362). New York: Plenum Press.

Anderson, D. R., & Lorch, E. P. (1983). Looking at television: Action or reaction? In J. Bryant & D. R. Anderson (Eds.), Understanding children's television (pp. 1-33). New York: Academic Press.

Andrews, M. L., & Madeira, S. S. (1977). The assessment of pitch discrimination ability in children. Journal of Speech and Hearing Disorders, 42, 279-286.

Barlow, N. L. B., Auslander, M. C., Rines, D., & Stelmachowicz, P. G. (1988). Probe-tube microphone measures in hearing-impaired children and adults. Ear and Hearing, 9, 243-247.

Beauchaine, K. L., Barlow, N. L. N., & Stelmachowicz, P. G. (1990). Special considerations in amplification for young children. Asha, 32, 44-46, 51.

Berger, K. W. (1976). Prescription of hearing aids: A rationale. Journal of the American Auditory Society, 2, 71-78.

Bonfils, P., Avan, P., Francois, M., Marie, P., Trotoux, J., & Narcy, P. (1990). Clinical significance of otoacoustic emission: A perspective. Ear and Hearing, 11, 155-158.

Boothroyd, A. (1967). The discrimination by partially hearing children of frequency distorted speech. International Audiology, 6, 136-145.

Boothroyd, A. (1968a). Developments in speech audiometry. British Journal of Audiology (Formerly Sound), 2, 3-10.

Boothroyd, A. (1968b). Statistical theory of the speech discrimination score. Journal of the Acoustical Society of America, 43, 362-267.

Boothroyd, A. (1970). Developmental factors in speech recognition. International Audiology, 9, 30-38.

Box, J. E. P. (1957). Evolutionary operation: A method for increasing industrial productivity. Applied Statistics, 6, 81-101.

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. Nature, 232, 456-458.

Byrne, D., & Cotton, S. (1988). Evaluation of the National Acoustic Laboratories' new hearing aid selection procedure. Journal of Speech and Hearing Research, 31, 178-186.

Byrne, D., & Dillon, H. (1986). The National Acoustic Laboratories' new hearing aid selection procedure. Ear and Hearing, 7, 257-265.

Byrne, D., & Tonisson, W. (1976). Selecting the gain of hearing aids for persons with sensorineural hearing impairments. Scandinavian Audiology, 5, 51-59.

Carhart, R. (1946). Tests for selection of hearing aids. Laryngoscope, 56, 780-794.

Carrow, E. (1973). Test for Auditory Comprehension of Language. Austin, TX: Learning Concepts.

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. 64-110). New York: Academic Press.

Clark, H. H. (1969). Linguistic processes in deductive reasoning. Psychological Reviews, 76, 387-404.

Collins, M. J., & Levitt, H. (1980). Comparison of methods for predicting optimum functional gain. In G. A. Studebaker & I. Hochberg (Eds.), Acoustical factors affecting hearing aid performance (pp. 341-354). Baltimore: University Park Press.

Cooper, R. L. (1967). The ability of deaf and hearing children to apply morphological rules. Journal of Speech and Hearing Research, 10, 77-86.

Cox, R. M. (1983). Using ULCL measures to find frequency/gain and SSPL 90. Hearing Instruments, 34, 17-21.

Davis, J. M. (1977). Reliability of hearing-impaired children's responses to oral and total presentations of the Test of Auditory Comprehension of Language. Journal of Speech and Hearing Disorders, 42, 520-527.

Donaldson, M., & Balfour, G. (1968). Less is more: A study of language and comprehension in children. British Journal of Psychology, 59, 461-472.

Donaldson, M., & Wales, R. (1970). On the acquisition of some relational terms. In J. R. Hayes (Ed.), Cognition and the development of language (pp. 235-238). New York: Wiley.

Ehri, L. C. (1976). Comprehension and production of adjectives and seriation. Journal of Child Language, 3, 369-384.

Elliott, L. L., & Katz, D. R. (1980a). Children's pure-tone detection. Journal of the Acoustical Society of America, 67, 343-344.

Elliott, L. L., & Katz, D.R. (1980b). Development of a new children's test of speech discrimination. St. Louis: Auditec.

Elliott, L. L., & Katz, D. R. (1980c). Northwestern University Children's Perception of Speech (NU-CHIPS): Technical Manual. St. Louis: Auditec.

Erber, N. P. (1973). Body-baffle effects and real-ear effects in the selection of hearing aids for deaf children. Journal of Speech and Hearing Disorders, 38, 224-231.

Erber, N. P. (1980). Use of the Auditory Numbers Test to evaluate speech perception abilities of hearing-impaired children. Journal of Speech and Hearing Disorders, 45, 527-532.

Finitzo-Hieber, T., Gerling, I. J., Matkin, N. D., & Cherow-Skalka, E. (1980). A sound effects recognition test for the pediatric audiological evaluation. Ear and Hearing, 1, 271-276,

Fior, R. (1972). Physiological maturation of auditory function between 3 and 13 years of age. Audiology, 11, 317-321.

Furth, H. G. (1964). Research with the deaf: Implications for language and cognition. The Psychological Bulletin, 62, 145-164.

Gabrielsson, A., Schenkman, B. N., & Hagerman, B. (1988). The effects of different frequency responses on sound quality judgments and speech intelligibility. Journal of Speech and Hearing Research, 31, 166-177.

Gabrielsson, A., & Sjogren, H. (1979). Perceived sound quality of sound-reproducing systems. Journal of the Acoustical Society of America, 65, 119-1033.

Gengel, R. W., Pascoe, D., & Shore, I. (1971). A frequency-response procedure for evaluating and selecting hearing aids for hearing-impaired children. Journal of Speech and Hearing Disorders, 36, 341-352.

Ginsburg, H., & Opper, S. (1969) Piaget's theory of intellectual development. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Gitterman, D., & Johnston, J. (1983). Talking about comparisons: A study of young children's comparative adjective usage. Journal of Child Language, 10, 605-621.

Goldman, R., & Fristoe, M. (1972). Goldman-Fristoe Test of Articulation. Circle Pines, MN: American Guidance Service, Inc. (AGS).

Gordon-Salant, S. (1984). Effects of reducing low-frequency amplification on consonant perception in quiet and noise. Journal of Speech and Hearing Research, 27, 483-493.

Green, R., Day, S., & Bamford, J. (1989). A comparative evaluation of four hearing-aid selection procedures. II--Quality judgements as measures of benefit. British Journal of Audiology, 23, 201-206.

Hall, J. W., & Ruth, R. A. (1985). Acoustic reflexes and auditory evoked responses in hearing aid evaluation. Seminars in Hearing, 6, 251-177.

Harris, R. W., & Goldstein, D. P. (1979). Effects of room reverberation upon hearing aid quality judgments. Audiology, 18, 253-262.

Haskins, H. A. (1949). A phonetically balanced test of speech discrimination for children. Unpublished master's thesis. Northwestern University, Evanston, IL.

Hawkins, D. B. (1985). Reflections on amplification: Validation of performance. Journal of the Academy of Rehabilitative Audiology, 18, 42-54.

Hawkins, D. B., Montgomery, A. A., Prosek, R. A., & Walden, B. E. (1987). Examination of two issues concerning functional gain measurements. Journal of Speech and Hearing Disorders, 52, 56-63.

Hawkins, D. B., Walden, B. E., Montgomery, A., & Prosek, R. A. (1987). Description and validation of an LDL procedure designed to select SSPL90. Ear and Hearing, 8, 162-169.

Hecox, K. E., & Punch J. L. (1988). The impact of digital technology on the selection and fitting of hearing aids. American Journal of Otology, 9, 77-86.

Inhelder, B., & Piaget, J. (1964). The early growth of logic in the child. New York: Humanities Press.

Jerger, S. (1984). Speech audiometry. In J. Jerger (Ed.) Pediatric audiology (pp. 71-93). San Diego: College-Hill Press, Inc..

Kawell, M. E., Kopun, J. G., & Stelmachowicz, P. G. (1988). Loudness discomfort levels in children. Ear and Hearing, 9, 133-136.

Lawson, G. D., & Chial, M. R. (1982). Magnitude estimation of degraded speech quality by normal- and impaired-hearing listeners. Journal of the Acoustical Society of America, 72, 1781-1787.

Layton, T., & Stick, S. (1979). Comprehension and production of comparatives and superlatives. Journal of Child Language, 6, 511-527.

Leijon, A., Eriksson-Mangold, M., & Bech-Karlsen, A. (1984). Preferred hearing aid gain and bass-cut in relation to prescriptive fitting. *Scandinavian Audiology*, *13*, 157-161.

Levitt, H. (1978a). Methods for the evaluation of hearing aids. In C. V. Ludvigsen & J. Barfod (Eds.), *Sensorineural hearing impairment and hearing aids*. *Scandinavian Audiology Supplement*, *6*, 199-240.

Levitt, H. (1978b). Adaptive testing in audiology. In C. V. Ludvigsen & J. Barfod (Eds.), *Sensorineural hearing impairment and hearing aids*. *Scandinavian Audiology Supplement*, *6*, 241-291.

Levitt, H., Neuman, A., & Toraskar, J. (1987). Orthogonal polynomial compression amplification for the hearing impaired. In R. D. Steel & W. Gerrey (Eds.), *Proceedings of the 10th Annual Conference on Rehabilitation Technology* (pp. 410-412). Washington D.C.: RESNA.

Levitt, H., Sullivan, J. A., Neuman, A. C., & Rubin-Spitz, J. A. (1987). Experiments with a programmable master hearing aid. *Journal of Rehabilitation Research and Development*, *24*, 29-54.

Libby, E. R. (1983). State-of-the-art of hearing aid selection procedures. *Hearing Instruments*, *36*, 30-38, 62.

Ling, D., & Ling, A. (1978). *Aural habilitation*. Washington D. C.: A. G. Bell Association for the Deaf.

Lippmann, O. (1971). Vision screening of young children. *American Journal of Public Health*, *61*, 1586-1601.

Lippmann, R. P., Braida, L. D., & Durlach, N. I. (1981). Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss. *Journal of the Acoustical Society of America*, *69*, 524-534.

Lybarger, S. F. (1978). Selective amplification--A review and evaluation. *Journal of the American Audiology Society*, *3*, 258-266.

Macpherson, B. J., Elfenbein, J. L., Schum, R. L., & Bentler, R. A. (1989). A procedure for obtaining thresholds of discomfort in young children. Paper presented at the American Speech-Language-Hearing Association 1989 Annual Convention. St. Louis, Missouri.

Matkin, N. D. (1986). Hearing aids for children. In W. R. Hodgson (Ed.), Hearing aid assessment and use in audiologic habilitation (3rd ed., pp. 170-190). Baltimore: Williams & Wilkins.

Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. Ear and Hearing, 3, 301-308.

McCandless, G., & Lyregaard, P. (1983). Prescription of gain/output (POGO) for hearing aids. Hearing Instruments, 34, 16-21.

Miller, P. H. (1979). Stimulus dimensions, problem solving, and Piaget. In G. A. Hale & M. Lewis (Eds.), Attention and cognitive development (pp. 97-118). New York: Plenum Press.

Montgomery, A. A., Schwartz, D. M., & Punch, J. L. (1982). Tournament strategies in hearing aid selection. Journal of Speech and Hearing Disorders, 47, 363-372.

Moore, J. M., & Wilson, W. R. (1978). Visual reinforcement audiometry (VRA) with infants. In S. E. Gerber & G. T. Mencher (Eds.), Early diagnosis of hearing loss (pp. 177-214), New York: Grune & Stratton.

Nelson, K., & Benedict, H. (1974). The comprehension of relative, absolute, and contrastive adjectives by young children. Journal of Psycholinguistic Research, 3, 333-342.

Neuman, A. C., Levitt, H., Mills, R., & Schwander, T. (1987). An evaluation of three adaptive hearing aid selection strategies. Journal of the Acoustical Society of America, 82, 1967-1976.

Nicolosi, L., Harryman, E., & Kresheck, J. (1978). Terminology of communication disorders. Baltimore: Williams & Wilkins.

Nielsen, L. B. (1989, May). A computer controlled digital master hearing aid. Paper presented at the 1989 International Symposium on Circuits and Systems. Portland, Oregon.

Northern, J. L., & Downs, M. P. (1984). Amplification for hearing-impaired children. In Hearing in children (3rd ed., pp. 269-303). Baltimore: Williams & Wilkins.

Nozza, R. J., & Wilson, W. R. (1984). Masked and unmasked pure-tone thresholds of infants and adults: Development of auditory frequency selectivity and sensitivity. Journal of Speech and Hearing Research, 27, 613-622.

Pascoe, D. P. (1975). Frequency responses of hearing aids and their effects on the speech perception of hearing impaired listeners. Annals of Otology, Rhinology and Laryngology, 84 (Suppl. 23), 3-40.

Pascoe, D. P. (1978). An approach to hearing aid selection. Hearing Instruments, 29 (6), 12-16, 36.

Pascoe, D. P. (1985). Hearing aid evaluation. In J. Katz (Ed.), Handbook of clinical audiology (3rd ed., pp. 936-948). Baltimore: Williams & Wilkins.

Pezdek, K., & Hartmann, E. F. (1983). Children's television viewing: Attention and comprehension of auditory versus visual information. Child Development, 54, 1015-1023.

Pezdek, K., & Stevens, E. (1984). Children's memory for auditory and visual information on television. Developmental Psychology, 20, 212-218.

Piaget, J. (1967). Six psychological studies. New York: Random House.

Pluvinage, V., & Benson, D. (1988). New dimensions in diagnostics and fitting. Hearing Instruments, 39 (8), 28-30, 39.

Punch, J. (1978). Quality judgments of hearing aid-processed speech and music by normal and otopathologic listeners. Journal of the American Audiology Society, 3, 179-188.

Punch, J. L., & Beck, E. L. (1980). Low-frequency response of hearing aids and judgments of aided speech quality. Journal of Speech and Hearing Disorders, 45, 325-335.

Punch, J. L., & Beck, L. B. (1986). Relative effects of low-frequency amplification on syllable recognition and speech quality. Ear and Hearing, 7, 57-62.

Punch, J. L., & Howard, M. T. (1978). Listener-assessed intelligibility of hearing aid-processed speech. *Journal of the American Auditory Society*, *4*, 69-76.

Punch, J. L., Montgomery, A. A., Schwartz, D. M., Walden, B. E., Prosek, R. A., & Howard, M. T. (1980). *Journal of the Acoustical Society of America*, *68*, 458-466.

Punch, J. L., & Parker, C. A. (1981). Pairwise listener preferences in hearing aid evaluation. *Journal of Speech and Hearing Research*, *24*, 366-374.

Raffin, M. J. M., Davis, J. M., & Gilman, L. A. (1978). Comprehension of inflectional morphemes by deaf children exposed to a visual English sign system. *Journal of Speech and Hearing Research*, *21*, 387-400.

Resnick, S. B. (1984). Speech recognition testing: Nonconventional testing techniques and application to children. In E. Elkins (Ed.), *Speech recognition by the hearing impaired*. *ASHA Reports*, *14*, 65-69.

Riley, C. A., & Trabasso, T. (1974). Comparatives, logical structures, and encoding in a transitive inference task. *Journal of Experimental Child Psychology*, *17*, 187-203.

Rintelmann, W. F., & Bess, F. H. (1988). High-level amplification and potential hearing loss in children. In F. H. Bess (Ed.), *Hearing impairment in children* (pp. 278-309). Parkton, MD: York Press, Inc.

Ross, M., & Lerman, J. W. (1970). A picture identification test for hearing-impaired children. *Journal of Speech and Hearing Research*, *13*, 44-53.

Ross, M., & Seewald, R. C. (1988). Hearing aid selection and evaluation with young deaf children. In F. H. Bess (Ed.), *Hearing impairment in children* (pp. 190-213). Parkton, MD: York Press, Inc.

Ross, M., & Tomassetti, C. (1980). Selecting amplification characteristics for young hearing-impaired children. In M. C. Pollack (Ed.), *Amplification for the hearing-impaired*, (2nd ed., pp. 213-253). New York: Grune & Stratton, Inc.

Schwartz, D. M., & Larson, V. D. (1977a). A comparison of three hearing aid evaluation procedures for young children. *Archives of Otolaryngology*, *103*, 401-406.

Schwartz, D. M., & Larson, V. D. (1977b). Hearing aid selection and evaluation procedures in children. In F. H. Bess (Ed.), Childhood deafness: Causation, assessment, and management (pp 217-233). New York: Grune & Stratton.

Seewald, R. C., & Ross, M. (1988). Amplification for young hearing impaired children. In M. Pollack (Ed.), Amplification for the hearing impaired (pp. 213-267). Orlando, FL: Grune & Stratton.

Seewald, R. C., Ross, M., & Spiro, M. K. (1985). Selecting amplification characteristics for young hearing-impaired children. Ear and Hearing, 6, 48-53.

Seewald, R. C., Ross, M., & Stelmachowicz, P. G. (1987). Selecting and verifying hearing aid performance characteristics for young children. Journal of the Academy of Rehabilitative Audiology, 20, 25-37.

Shapiro, I. (1976). Hearing aid fitting by prescription. Audiology, 15, 163-173.

Shaw, E. A. G. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. Journal of the Acoustical Society of America, 56, 1848-1861.

Shaw, E. A. G., & Vaillancourt, M. M. (1985). Transformation of sound-pressure level from the free field to the eardrum presented in numerical form. Journal of the Acoustical Society of America, 78, 1120-1123

Shore, I., Bilger, R. C., & Hirsh, I. J. (1960). Hearing aid evaluation: Reliability of repeated measurements. Journal of Speech and Hearing Disorders, 25, 152-170.

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

Siegenthaler, B. M. (1969). Maturation of auditory abilities in children. International Audiology, 8, 59-71.

Skinner, M. W. (1980). Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation. Journal of the Acoustical Society of America, 67, 306-317.

Skinner, M. W., Pascoe, D. P., Miller, J. D., & Popelka, G. R. (1982). Measurements to determine the optimal placement of speech energy within the listener's auditory area: A basis for selecting amplification characteristics. In G. A. Studebaker, & F. H. Bess (Eds.), The Vanderbilt hearing-aid report: State of the art--research needs (pp. 161-169). Upper Darby, PA: Monographs in Contemporary Audiology.

Streng, A. H., Kretschmer, R. R., & Kretschmer, L. W. (1978). Language, learning, and deafness. New York: Grune & Stratton, Inc.

Stuart, A., Durieux-Smith, A., & Stenstrom, R. (1990). Critical differences in aided sound field thresholds in children. Journal of Speech and Hearing Research, 33, 612-615.

Studebaker, G. A. (1982). Hearing aid selection: An overview. In G. A. Studebaker & F. H. Bess (Eds.), The Vanderbilt hearing-aid report: State of the art--research needs (pp. 147-155). Upper Darby, PA: Monographs in Contemporary Audiology.

Studebaker, G. A., Bisset, J. D., Van Ort, D. M., & Hoffnung, S. (1982). Paired comparison judgments of relative intelligibility in noise. Journal of the Acoustical Society of America, 72, 80-92.

Studebaker, G. A., & Sherbecoe, R. L. (1988). Magnitude estimations of the intelligibility and quality of speech in noise. Ear and Hearing, 9, 259-267.

Sullivan, J. A., Levitt, H., Hwang, J-Y., & Hennessey, A-M. (1988). An experimental comparison of four hearing aid prescription methods. Ear and Hearing, 9, 22-32.

Tecca, J. E., & Goldstein, D. P. (1984). Effect of low-frequency hearing aid response on four measures of speech perception. Ear and Hearing, 5, 22-29.

Thornton, A. R., & Raffin, M. J. M. (1978). Speech-discrimination scores modeled as a binomial variable. Journal of Speech and Hearing Research, 21, 507-518.

Thompson, G., & Lassman, F. (1969). Relationship of auditory distortion test results to speech discrimination through flat vs selective amplifying systems. Journal of Speech and Hearing Research, 12, 594-606.

Thompson, G., & Lassman, F. (1970). Listener preference for selective vs flat amplification for a high-frequency hearing-loss population. Journal of Speech and Hearing Research, 13, 667-672.

Walden, B. E., Schwartz, D. M., Williams, D. L., Holum-Hardegen, L. L., & Crowley, J. M. (1983). Test of the assumptions underlying comparative hearing aid evaluations. Journal of Speech and Hearing Disorders, 48, 264-273.

Wannemacher, J. T., & Ryan, M. L. (1978). "Less" is not "more": A study of children's comprehension of "less" in various task contexts. Child Development, 49, 660-668.

Watson, N. A., & Knudsen, V. O. (1940). Selective amplification in hearing aids. Journal of the Acoustical Society of America, 2, 406-419.

Witter, H. L., & Goldstein, D. P. (1971). Quality judgments of hearing aid transduced speech. Journal of Speech and Hearing Research, 14, 312-322.

Wood, B. S. (1976). Children and communication: Verbal and nonverbal language development. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Zerlin, S. (1962). A new approach to hearing-aid selection. Journal of Speech and Hearing Research, 5, 370-376,