A Modification of Play Audiometry to Assess Speech Discrimination Ability in Severe-Profoundly Deaf 2- to 4-Year-Old Children.

P.W. Dawson, P.E. Nott, G.M. Clark, and R.S.C. Cowan

Objective: The aim was to develop an assessment procedure that was independent of language and speech production ability, to test speech feature discrimination in severe-profoundly deaf children 2 to 4 yr of age.

Design: The procedure being trialed was adapted from existing procedures. The child was required to respond with a game-like motor response to a "change" in a speech stimulus that was being presented repeatedly through a speaker. The change occurred at randomly determined times, and false alarm responses were measured during the waiting periods (while the child waited for the change). Two- to four-yr-old normally hearing children and hearing-impaired children using hearing aids and a group of 4-yr-old hearing-impaired children using cochlear implants were assessed on the task.

Results: More than 82% of the 3- and 4-yr-old normally hearing and hearing-impaired children were able to complete the testing for the eight speech sound contrasts within three 20 minute sessions. Fifty percent of the 2-yr-old normally hearing and hearing-impaired children were able to condition and complete the task. All of the normally hearing children who completed the task successfully discriminated all speech sound contrasts. The performance of the hearing-impaired children using hearing aids was influenced by the degree of hearing loss and the type of speech contrast being tested. Similarly, the average performance of the children using cochlear implants was better for easier contrasts such as /ba/bi/ with contrasting vowel formant cues.

Conclusions: This procedure has potential for use as a reliable clinical and research tool for assessing the development of auditory discrimination ability in 2- to 4-yr-old severe-profoundly deaf children.

(Ear & Hearing 1998;19:371-384)

An increasing number of very young children are being implanted with multichannel cochlear implants, yet there is little known about the development of their ability to utilize the auditory cues coded by the implant. Conventional speech perception tests are often inappropriate for evaluating the auditory capacity of young, severe-profoundly deaf children with limited language. It is important to assess this age group with a speech perception test that is analytical and able to measure the child's pure auditory capacity without the confounding influence of the child's ability to use linguistic information. Other factors, such as the young child's cognitive status, motivation, and ability to sustain attention, also must be given serious attention.

Boothroyd (1996) has developed an imitative version (IMSPAC) of his Speech Perception Contrast Test (SPAC; Boothroyd, 1984) that satisfies some of these criteria. The child is asked to imitate a syllable such as /uk/ using auditory-visual input (and print if appropriate) and then using auditory-alone input. A listener then decides whether the child was imitating /uk/, /ug/, /ik/, or /ig/, thereby accurately conveying the features of vowel place and final consonant voicing. The test aims to provide information on auditory capacity compared with an estimate of best production (i.e., the auditory-visual score). Results from 97 hearing-impaired children aged 3 to 15 yr revealed that the auditory-alone scores were a measure of auditory capacity and were not influenced by maturation and learning. Although some very young children were able to do the task, Boothroyd (1984) emphasizes that the test is not suitable for prelingually deaf children with poor speech production skills.

Like Boothroyd (1984), many other researchers have recognized the need to develop speech perception assessments that are more suitable for the severe-profoundly deaf preschool child. Test development has been characterized by the use of smaller closed sets of alternatives, simpler vocabulary, and concrete objects rather than ambiguous pictures from which the child selects. The Low Verbal version of the Early Speech Perception test uses closed sets of four concrete objects and assesses recognition of words differing in syllable or segmental pattern (Moog & Geers, 1990). Similarly, one part of the Grammatical Analysis of Elicited Language-Pre-school Level (Moog, Kozak, & Geers, 1983) was...
modified by Robbins and Kirk (1996) to measure auditory-alone word recognition using closed sets of four concrete objects. Robbins and Kirk also reported on the development of the Mr. Potato Head Task, a modified open-set task for preschool children. This task requires the child to follow commands in assembling a Mr. Potato Head toy through listening alone. However, in each of the assessments described above, the child is required to have knowledge of specific vocabulary, which in many cases may need to be taught. Furthermore, there is an age effect on some of these tasks for normally hearing children in that older children perform better on average than younger children. For example, the mean open-set word and sentence comprehension scores on the Mr. Potato Head Task were about 60% for 2-yr-old normally hearing children compared with approximately 90% and 100% for 3- and 4-yr-old children, respectively.

The Screening Inventory of Perception Skills is a discrimination assessment using a go/no go paradigm, which does not require knowledge of certain vocabulary (Osberger et al., 1991). Real objects are used instead of pictures in a recent modification (Robbins & Kirk, 1996). For example, the child puts a fish in a bag when he or she hears the target word /fish/ but not when there is a foil such as /toothbrush/. Like the Low Verbal version of the Early Speech Perception task, this test has four levels, moving from syllable pattern discrimination to gross segmental discrimination. An age effect was also evident for this test in that normally hearing 2-yr-old children were found to perform less well than the 3- and 4-yr-old children. The children younger than 2.5 yr of age actually performed better when an alternative response mode of imitation was used. The task is appealing because it is quick to administer, but it provides only gross information about auditory capacity. More analytical information is essential to evaluate the efficacy of the coded speech information provided by multichannel cochlear implants.

Normally hearing infants have been assessed with an analytical, speech-feature discrimination task that is independent of language ability. The Visual Reinforcement Infant Speech Discrimination test (VRISD) is adapted from the operant, headturn procedure for infant acuity testing (Moore, Wilson, & Thompson, 1977). In the VRISD task the infant learns to turn to a change (e.g., /pa/) in a repeating speech stimulus (e.g., /ba/) and is rewarded by an illuminated puppet. A number of speech contrasts or features are assessed. No-change trials are included to assess the tendency of the infant to make random head turns. This procedure has been used to evaluate the development of speech discrimination (Eilers, Wilson, & Moore, 1977) and the perceptual constancy of vowel and consonant categories in normally hearing infants (Kuhl, 1979, 1980). Infants are not, however, always able to condition to the task or able to complete the task (Aslin, Pisoni, Hennessy, & Perey, 1981; Nozza, Miller, & Bond, 1991). The VRISD procedure has been administered to preschool children using different stimuli and a different response mode. The ability of 11 normally hearing 4-yr-old children to discriminate real word pairs using the VRISD procedure with a button press response was compared with their ability to discriminate the word pairs using a conventional picture pointing task (Menary, Trehub, & McNutt, 1982). Similar levels of performance were reported for both response modes.

The “change/no change” task also was derived from the VRISD task, using speech stimuli similar to those in the VRISD task but a different response mode (Sussman & Carney, 1989). Either a motor response (e.g., button push) or a same/different verbal response was required. The change/no change task was used to assess speech discrimination, selective adaptation, and/or labeling in normally hearing children as young as 5 yr of age (Sussman, 1993; Sussman & Carney, 1989) and discrimination ability in severely-profoundly deaf children as young as 4 yr of age (Carney, Osberger, Carney, Robbins, Renshaw, & Miyamoto, 1993; Eilers, Cobo-Lewis, Vergara, Oller, & Friedman, 1996; Osberger, Miyamoto, Zimmerman-Phillips, Kemink, Stroer, Firszt, & Novak, 1991; Osberger, Robbins, Miyamoto, Berry, Myres, Kessler, & Pope, 1991).

In a comparison of children’s speech discrimination with cochlear implants and tactile aids, Carney et al. (1993) assessed nine speech contrasts with 15 trials per contrast; 10 change trials and five no-change trials. Children were required to respond during the change trial and not during the no-change trial. In a change trial there were five repetitions of the background stimuli followed by five repetitions of a change stimulus, whereas in the no-change trial there were 10 repetitions of the background stimulus. Results were analyzed in terms of the number of correct responses, consisting of hits (i.e., responses during change trials) and correct rejections (i.e., no responses during no-change trials). A measure of the sensitivity of the discrimination responses was obtained by calculating d’. The test was found to be sensitive to differences in performance between the groups of hearing-impaired children using the different devices. Some of the younger hearing-impaired children in this study and in the study by Osberger, Robbins, Miyamoto, Berry, Myres, Kessler, and Pope (1991),
TABLE 1. Numbers of children who passed the conditioning criterion and numbers of children who completed all eight speech contrasts. The means and standard deviations of age are given for each of the 2-, 3-, and 4-yr-old subgroups tested.

<table>
<thead>
<tr>
<th>Group</th>
<th>Age (yr)</th>
<th>Number of Subjects Tested</th>
<th>Number Conditioned</th>
<th>Number Completed Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal hearers</td>
<td>2.45 (0.21)</td>
<td>10</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3.56 (0.33)</td>
<td>10</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>4.60 (0.30)</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Hearing aid users</td>
<td>2.69 (0.21)</td>
<td>12</td>
<td>10 (6)</td>
<td>6 (9)</td>
</tr>
<tr>
<td></td>
<td>3.67 (0.23)</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>4.54 (0.31)</td>
<td>12</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Cochlear implant users</td>
<td>3.25*</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4.55 (0.28)</td>
<td>11</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

*This is the age of the single subject in the 3-yr-old age group.

However, were either unable to condition to the test or unable to complete the test.

The VRISD task has been modified successfully for use with severe-profoundly deaf school-aged children, but difficulties have arisen in assessing the 4-yr-old population. Can it be modified further for use with the challenging 2- to 4-yr-old population? This study aimed to develop a Speech Feature Test (SFT) that shares some similarities with the VRISD task and the change/no change task but that focuses on incorporating aspects of the familiar “play audiometry” procedure used in hearing acuity testing of very young children (Northern & Downs, 1991).

In contrast to the change/no change procedure, the background stimulus is presented soon after the child enters the test room in the SFT. This is designed to prevent the child from responding to the onset of the background stimulus. In a pilot study by King and Rance (Reference Note 1) the background stimulus was introduced as soon as the child was in a ready state of listening; that is, had placed a toy next to his or her ear and was attentive. The 3-yr-old children in the study confused the task with hearing acuity testing and responded to the onset of the background stimulus. To reduce overall test time and task complexity, the current procedure does not include separate change and no-change trials. Instead, the child is conditioned to respond to a change in the stimulus, which occurs at a randomly determined time after the child enters a ready state. In the change/no change procedure, each trial is of the same duration but can finish with or without a change occurring. It is likely that the very young child will be confused if a trial has a set length and can end sometimes without the motor response being required. The current procedure was developed deliberately to be more akin to play audiometry, which employs variable waiting times before the presentation of audible tones. A trial in the present procedure simply refers to the occurrence of a change stimulus, and false positive responses are recorded during the waiting times. There are distinct advantages to using a technique with which young hearing-impaired children are familiar.

This study was specifically designed to determine:

1. Whether normally hearing 2- to 4-yr-old children could perform the SFT procedure.
2. Whether the SFT discrimination performance of children using hearing aids is influenced by the children’s age at testing.
3. Whether performance on the task is related to degree of hearing loss.
4. Whether the discrimination performance of children using hearing aids and children using cochlear implants varies as a function of the particular speech contrast being assessed.
5. Whether performance on the task is reliable from one test occasion to another.

**Method**

**Subjects**

A total of 78 children residing in metropolitan Melbourne were assessed using the SFT. They included 30 normal hearers, 36 hearing aid wearers, and 12 cochlear implant users (see Table 1).

**Normal Hearers** • Table 1 shows that 30 normally hearing children were divided into three equal groups of 2, 3, and 4 yr olds. All children were volunteers attending an integrated child care and kindergarten program. None of the children demonstrated observable developmental problems. All demonstrated normal tympanometry, which was defined as appropriate compliance peaks for each ear at a middle ear pressure no more negative than −150 mm H₂O.

**Hearing Aid Wearers** • Thirty-six children using hearing aids were assigned to one of three hearing loss groups based on their unaided pure-tone average (PTA) hearing losses (three frequency average at 0.5, 1, 2 kHz in the better ear). The PTAs ranged from 31 to 59 dB HL, 60 to 85 dB HL, and 90 to 113 dB HL for the mild to moderate, severe, and pro-
<table>
<thead>
<tr>
<th>Speech Contrast</th>
<th>Syllable Duration</th>
<th>F0 (Start)</th>
<th>F0 (End)</th>
<th>Consonant Duration (msec)</th>
<th>F1</th>
<th>F2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td>506</td>
<td>240</td>
<td>234</td>
<td>8</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>b-</td>
<td>253</td>
<td>242</td>
<td>240</td>
<td>9</td>
<td>800</td>
<td>1400</td>
</tr>
<tr>
<td>F0 steady/F0 falling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td>506</td>
<td>240</td>
<td>234</td>
<td>8</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>ba-</td>
<td>518</td>
<td>360</td>
<td>240</td>
<td>9</td>
<td>760</td>
<td>1360</td>
</tr>
<tr>
<td>Vowel height</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td>506</td>
<td>240</td>
<td>234</td>
<td>8</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>bi</td>
<td>503</td>
<td>242</td>
<td>240</td>
<td>6</td>
<td>400</td>
<td>2360</td>
</tr>
<tr>
<td>Vowel place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bi</td>
<td>503</td>
<td>242</td>
<td>240</td>
<td>6</td>
<td>360</td>
<td>2360</td>
</tr>
<tr>
<td>bu</td>
<td>516</td>
<td>247</td>
<td>242</td>
<td>12</td>
<td>400</td>
<td>1560</td>
</tr>
<tr>
<td>Consonant manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td>508</td>
<td>240</td>
<td>234</td>
<td>18</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>ma</td>
<td>526</td>
<td>236</td>
<td>240</td>
<td>136</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>Consonant manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sa</td>
<td>525</td>
<td>247</td>
<td>240</td>
<td>144</td>
<td>760</td>
<td>1360</td>
</tr>
<tr>
<td>ta</td>
<td>514</td>
<td>242</td>
<td>240</td>
<td>80</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>Consonant voicing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td>506</td>
<td>240</td>
<td>234</td>
<td>8</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>pa</td>
<td>508</td>
<td>247</td>
<td>240</td>
<td>53</td>
<td>760</td>
<td>1360</td>
</tr>
<tr>
<td>Consonant place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba</td>
<td>506</td>
<td>240</td>
<td>234</td>
<td>8</td>
<td>800</td>
<td>1360</td>
</tr>
<tr>
<td>ga</td>
<td>503</td>
<td>237</td>
<td>235</td>
<td>9</td>
<td>760</td>
<td>1400</td>
</tr>
</tbody>
</table>

Note: The background syllable appears first in order in each pair.
The 1st and 2nd formants of the vowels are represented as F1 and F2.

found groups, respectively. There were 12 children in each hearing loss group, divided equally into three groups of 2, 3, and 4 yr olds. Their mean ages are shown in Table 1. Thirty-four of the children were congenitally deaf, and two acquired deafness at 10 mo of age. Etiology was unknown for 17 children, genetic for 11, cytomegalovirus for two, birth trauma for four, and meningitis for two. Twenty-nine children attended an integrated early intervention program, and the remaining four children was total communication in which Signed English was used. All hearing aids were checked on the day of testing for recommended volume setting and battery function, and a daily listening check was performed by teachers of the deaf. All children demonstrated normal tympanometry on the days of testing.

**Cochlear Implant Users** - The cochlear implant group comprised one 3 yr old and 11 4 yr olds using the 22-electrode cochlear implant with the Spectra speech processing strategy. All of the children were congenitally deaf and implanted between the ages of 1 yr 8 mo and 3 yr 10 mo with implant experience varying from 10 mo to 3 yr 3 mo. Etiology was cytomegalovirus for two of the children, charge syndrome for five, and meningitis for two. All children demonstrated normal tympanometry on the days of testing.

unknown for the remaining children. Ten of the children had 20 active electrodes and used common ground mode of stimulation in which the current flows from the active electrode to all other electrodes. One implant user had 18 active channels and used pseudo-monopolar stimulation in which the current flows from the electrode being stimulated to the most basal electrode. Another implant user had 12 active channels in a wide mode of stimulation, "bipolar + 4," in which electrodes are activated as pairs. The electrodes in each pair were separated by four electrodes. Seven children attended an integrated setting and five attended a segregated setting for hearing-impaired children. The mode of communication for nine of the children was oral and for the remaining three children was total communication.

**Stimuli**

A number of nonmeaningful speech contrasts were used, consisting of repeated CV syllables, one acting as a "background stimulus" the other as a "change" stimulus. Where possible the background stimulus remained the same, i.e., /hal/ to facilitate transfer of learning across speech contrasts. Eight speech contrasts were chosen to test a range of speech features, including suprasegmental and segmental features (see Table 2). Several of these have
been found to be sensitive to differences in performance amongst profoundly deaf children using a variety of devices (Carney et al., 1993; Osberger, Robbins, Miyamoto, Berry, Myres, Kessler, & Pope, 1991). The speech contrast \textipa{/ba/\textipa{mi}/\textipa{\gamma/}} used in the initial training phase maximized differences in speech feature information, with differences in vowel height, consonant manner, nasality, and intonation. Some speech contrasts similar to those in the test phase also were included in the set of training contrasts to facilitate transfer of learning during the test phase when necessary (see end Stimulus Presentation Format section). A 10 dB increase in intensity could be used for the change stimulus in any of the training contrasts.

Stimuli were produced by a female speaker and recorded directly onto a computer (16 bit PCM format) as individual digital audio files using a sampling rate of 22 kHz with a Pro Audio Spectrum 16 Sound Card and a Shure SM57 Dynamic microphone. Table 2 gives details of some acoustic properties of the test stimuli, which were measured on the Kay Spectrograph (Model 5500). Overall length in each of the manner contrasts was matched as closely as possible in a way that maintained the natural difference in the duration of the consonants by manipulating the length of the vowel but not to such a degree that a change was likely to be perceived on vowel length alone. Table 2 indicates that syllables within a contrast also were matched closely for fundamental frequency (F0) at the start and end of the syllables. All syllables were RMS equalized. The wave form envelope was manipulated at the end of stimuli by fading to achieve a natural vowel quality. An interval of 800 msec occurred between syllables during testing.

**Equipment**

Portable equipment was selected so that testing could be conducted at the children's educational setting. A portable computer (AcerNote 950) with a 16 bit built-in sound card was used to run a computer program designed to administer the test. Speech stimuli used in the test were presented by a self-amplified portable speaker (BOSE Lifestyle Powered Speaker System) with a wide frequency response.

Testing was performed in a quiet room with children seated at a table in front of the speaker. The speaker was mounted on a tripod positioned 1 meter from the child at the child's ear level. It was calibrated using a sound level meter (CEL-245 Digital Impulse) to 68 dBA at the ear level of the child. An assistant tester was seated alongside the child. The tester wore headphones (Sennheiser HD 435 Manhattan) and was seated at the computer, which was positioned at approximately 90 degrees to the child.

**Procedure**

Six trials were presented for each of eight speech contrasts with a total of 48 trials per subject, compared with 135 trials in the change/no change procedure (Carney et al., 1993). A score out of six is given for each speech contrast, which represents the number of correct responses for six change presentations or trials. False positive response rates (per second) are calculated by dividing the number of false positive responses that occur during the total waiting time over the six trials by the duration of the total waiting time in seconds (total time excluding the “hit” windows in which the change stimulus occurred). The probability that a false positive response will occur in a hit window (i.e., the child responds without actually perceiving the change) then is considered to be equal to the false positive rate multiplied by 4 sec, the length of the hit window. The simplifying assumption is made that the false positive rate is a constant.

The binomial distribution was used to determine an acceptable number of trials with reasonable power. The criterion for successfully “passing” a speech contrast depended on the probability of a false positive in the hit window. If this probability did not exceed 0.26, a score of ≥4/6 represented a “pass” and was significantly above “chance” performance at $p < 0.05$. If this probability was higher, 0.3 for example, a higher score than 4/6 was needed to achieve the 95% confidence limit. When the assumed probability of a child actually perceiving a change on an individual item is 0.75, the power or sensitivity of the test with a criterion score of ≥4/6 is reasonably high at 83%. That is, with this criterion, one could expect to miss only 17% of children who can actually detect a difference in the stimuli at this assumed level of discrimination.

The current procedure used a variety of motor responses (placing a toy in a bucket or down a ramp or candles in playdoh) in any one session of testing. Frequent changes of activity were employed with the 2-yr-old children to maintain interest and attention in the task. Northern and Downs (1991) note the importance of game changes in play audiometry. A consistent, salient behavioral response (e.g., verbal “ah” with smile or head turn to the computer) was accepted in 2- to 3-yr-old children who were tentative in performing the required motor response. Both the assistant and the tester needed to vote that a response had occurred for the response to be accepted.
which the assistant tester had his or her hand over formed the motor responses for the child. When contrast and testing began once the child was in a ready state. For hearing-impaired children the easier contrasts /ba/ba\xy, /ba/bp\ and /ba/bi/ were tested first in random order, followed by /ba/mal, and finally the remaining contrasts in random order. For normally hearing children testing began with the /ba/bi/ contrast because this was felt to be an easy contrast for infants tested using the VRISD task (Kuhl, 1979). This contrast was followed by a random ordering of the remaining contrasts. A protocol of feedback was adopted for the test phase based on that used in play audiometry. Positive verbal feedback was given to the child for responding during the change window. No feedback was given when the child failed to respond during the change window. If a false positive response occurred, the response was not accepted and the child was encouraged to wait. All responses were recorded on the computer by the tester, who wore headphones and listened to a complex wave file consisting of a combination of different speech contrast wave files. This was found to be an effective and comfortable masker of the change stimulus and reduced potential bias by the tester when recording responses. Responses were automatically scored as correct responses or false positives by the computer according to whether they occurred during a change window or a wait period, respectively.

A failed contrast (score <4/6) could be due to a loss of conditioning or to an inability to generalize the task across different contrasts. For children who failed a speech contrast at any time during the testing, a conditioning check was performed to ensure that the child had not failed because of a loss of conditioning. Northern and Downs (1991) suggest that several reconditioning periods may be necessary when testing young children using a play audiometry task. The check involved presenting a trial from a contrast previously “passed” easily (≥5/6 if possible). If the child responded correctly, the score was retained. If the child did not respond to the change stimulus, a training phase was introduced to reestablish conditioning. If conditioning was re-established, the failed speech contrast was repeated and the retest score accepted. The conditioning check also served to provide the child with some positive reinforcement. Similarly, Hazan, Fourcin, and Abberton (1991), in a study of severe-profoundly deaf older children’s ability to perceive and label initial voiced plosives along a six step continuum, presented endpoint stimuli now and then to prevent the children from becoming frustrated after hearing ambiguous stimuli.

Inability to generalize the task to “any” change occurring in a repeating background stimulus also could contribute to failure on a contrast. Transfer of learning can be assumed to occur in the initial.

**Stimulus Presentation Format** During the training phase the timing of presentation of the background and change stimuli were controlled completely by the tester to allow for maximum flexibility. Figure 1 provides an outline of the timing parameters for stimulus presentation in the test phase. The length of the wait periods were variable, being determined randomly by the computer; all change stimuli were presented for a duration of 4 sec.

As soon as the child entered a ready state in the test phase, a computer key was activated to begin a “wait” period, which was followed by the first trial or presentation of a change stimulus. If the child responded during the change period, the repeating background stimulus automatically returned and the procedure was repeated for subsequent trials. If the child failed to respond during the change window, the background stimulus returned for a minimum of 3 sec before a second change stimulus occurred. To avoid the young child waiting for excessively long periods of time, the total period of the Wait 1 + Change Stimulus + Wait 2 never exceeded 15 sec. If the child failed to respond to two consecutive change stimuli, testing was interrupted by the assistant to give the child verbal encouragement before continuing with subsequent trials. If a child lost readiness at any stage after testing had begun, that is, dropped the toy or talked, the trial was repeated but the original waiting time was included in calculating the false positive rate.

Each child was given a maximum of three 20 minute sessions to complete the eight speech contrasts. Each session began with a training phase in which the tester selected the training speech contrast /ba/mi\xy. The assistant tester “shaped” the child to wait and respond with a motor response to the change stimulus. Four “shapes” were provided in which the assistant tester had his or her hand over the child’s hand that was holding the toy and performed the motor responses for the child. When three consecutive independent responses were obtained for /ba/mi\xy without an intensity increase on mi\xy, the conditioning criterion was met and the test phase began.

In the test phase, the tester selected a speech contrast and testing began once the child was in a

<table>
<thead>
<tr>
<th>Background Stimulus</th>
<th>Wait 1</th>
<th>Change 1</th>
<th>Wait 2</th>
<th>Change 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ba……………….ma…….ba……………….ma…….ba.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Timing parameters for stimulus presentation in the test phase.
stages of testing when a cross-section of different types of contrasts are presented. For children who failed any of the first four contrasts, and were still found to be conditioned, additional training trials were provided to ensure that a lack of transfer of learning had not contributed to the failure. Four training trials were provided, using a training contrast that was similar to that failed in the test phase. The tester shaped the child's responses during these trials. The test contrast then was retested. An increase in score of ≥2 was seen as suggestive of lack of generalization, so the retest score was accepted in these cases.

Test-retest data were obtained for 12 children. Each of the these children was tested on two contrasts, and the contrasts were pseudo-randomly selected to ensure near-equal representation of all contrasts. Three children were selected randomly from each of the groups of 2-, 3-, and 4-yr-old hearing aid wearers and implant users.

**RESULTS**

**Attrition Rates**

The numbers of 2-, 3-, and 4-yr-old children (normal hearers and children using hearing aids or cochlear implants) who passed the conditioning criterion are shown in Table 1. It can be seen that 30% of the 2-yr-old normally hearing children and 17% of the 2-yr-old hearing aid users could not be conditioned. One 4-yr-old using a cochlear implant also could not be conditioned. Not all of the children who were able to be conditioned could complete the task within three 20 minute sessions. Table 1 indicates that two of the 2-yr-old normally hearing and four of the 2-yr-old hearing-impaired children were unable to complete all of the eight speech sound contrasts. Also, the 3-yr-old implanted child and three 4-yr-old hearing aid users could not complete the task.

**Ease and Efficiency of Testing**

Figure 2 shows that in the normally hearing group the older children generally took only one or two sessions to complete the task, whereas the 2-yr-old children took two or three sessions on average. This effect of age was not evident for the hearing aid users, who, like the 4-yr-old implant users, usually took 2 or 3 sessions to complete the testing. The lower graph in Figure 2 reveals that the hearing aid users with a mild to moderate degree of hearing loss took less time on average to complete the testing than did the children with a greater hearing impairment.

Of those who completed the task, some had evidence of difficulty with transfer of learning (see Method section). Two of the five 2-yr-old normally hearing children and one of the six 2-yr-old hearing-impaired children needed help to facilitate generalization. Four of the nine normally hearing 3-yr-old children and four of the 12 hearing-impaired 3 yr olds also showed difficulty. In contrast, no normally hearing 4-yr-old child revealed a difficulty with transfer of learning, but one of the nine 4-yr-old aid users and three of the 10 4-yr-old implant users experienced difficulty. Transfer difficulty usually occurred for only one to two contrasts per child. In all cases the children scored ≤2/6 on the contrast initially but after "shaping" on a similar training contrast, they scored ≥5/6.

**Speech Discrimination Performance**

Statistical analyses were conducted on the results for those children who completed the eight contrasts.

**Normally Hearing Children** • A 2-way analysis of variance (ANOVA) with ranked data and one inde-
Hearing Aid Users. A 2-way ANOVA using ranked data and one independent factor (age) and one repeated measures factor (type of contrast) revealed that contrast type was a significant main effect \((F(7, 168) = 5.74; p < 0.0001)\). Figure 3 shows that the mean discrimination scores for the contrasts /balbr/, /balba/, and /balbi/ were higher than those for the other contrasts. A Student-Newman-Keuls multiple comparison test indicated significant differences between performance for each of these three contrasts and the other five contrasts. Mean scores for the contrasts ranged from 5.52 for /ba/sa/ to 5.89 for /ba/ga/.

Figure 3. Hearing aid users' mean discrimination scores (maximum = 6) and standard errors for each speech contrast as a function of age group. Speech contrasts are represented here by the “change” stimulus only.

There was a significant main effect for hearing loss in a 2-way ANOVA using ranked data and one independent factor (hearing loss) and one repeated measures factor (type of contrast) \((F(2, 24) = 25.65; p < 0.0001)\) (see Fig. 4). There were significant differences between all combinations of the three hearing loss groups. Means averaged across contrasts were 5.84, 4.82, and 3.57 for the mild to moderate, severe, and profound groups, respectively. The mean PTAs and SDs of the means for the mild to moderate, severe, and profound groups were 43.9 (8.08), 71.1 (8.11), and 102.3 (8.41). A significant interaction occurred between hearing loss group and type of contrast \((F(14, 168) = 2.28; p = 0.0071)\). The means and standard errors in Figure 4 help reveal that significant differences between hearing loss groups were present only for certain speech contrasts: for /bi/bu/, /ba/ma/, /ba/pa/, /ta/ssa/, and /ba/ga/. In all cases except /ba/ga/, significant mean differences occurred only between the mild to moderate and profound hearing loss groups. For /ba/ga/, a significant difference also occurred between the mild to moderate and severe groups. The effect of contrast type also was present only at certain levels of the hearing loss factor. For the mild to moderate group, there were no significant differences in performance between pairs of contrasts. Although differences in the severely hearing-impaired children's mean performance for different contrasts can be seen in Figure 4, post hoc analyses revealed no significant differences between pairs of
TABLE 3. Probabilities of a false positive response occurring in the “hit” window.

<table>
<thead>
<tr>
<th></th>
<th>Normal Hearers</th>
<th>Hearing Aid Users</th>
<th>Implant Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.027</td>
<td>0.029</td>
<td>0.048</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.058</td>
<td>0.057</td>
<td>0.070</td>
</tr>
<tr>
<td>Range</td>
<td>0.000–0.236</td>
<td>0.000–0.296</td>
<td>0.000–0.332</td>
</tr>
<tr>
<td>Median</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

be significantly different /ba/bi/ versus /ta/sa/ and /ba/bi/ versus /ba/pa/.

Test-Retest Reliability

A Pearson Product Moment Correlation was computed and showed a highly significant correlation between the test and retest assessments that involved 12 of the children ($R = 0.916; p < 0.0001$). A high correlation can be obtained without scores being almost the same. For example, a learning effect may result in the scores differing by a constant amount. Therefore, another measure of reliability was computed: the stability of the scores, measured as the SD of the differences between the test and retest scores. The mean difference score was 0.21, and the SD of the difference scores was 0.98. Therefore, in 95% of cases, the retest score would be expected to deviate from the original test score by no more than two SDs (approximately two points out of six for a single subtest).

False Positive Responses

For the normally hearing children and the hearing aid users, there were no significant differences in the false positive rates across age or contrasts. For the children using cochlear implants, there were no significant differences in the false positive rates across contrasts. Table 3 shows that the mean probabilities of a false positive response occurring in the hit window for the normally hearing children and for the children using hearing aids and cochlear implants were low. Ranges, SDs, and medians also are given.

DISCUSSION

Attrition Rates

Given that the procedure is similar to play audiometry in many respects, it was anticipated that some children, especially those in the 2-yr-old group, would fail to condition. The percentages of 2-yr-old normally hearing and hearing-impaired children in the present study who failed to condition were consistent with those found for play audiometry.
Thompson, Thompson, and Vethivelu (1989) reported that 32% of younger 2-yr-old normally hearing children could not be conditioned to play audiometry. A greater percentage of normally hearing than hearing-impaired 2-yr-old children failed to condition in the current study. It should be remembered, however, that the mean age of the 12 2-yr-old hearing aid users was greater than that of the group of 10 normally hearing 2-yr-old children assessed. The three normally hearing children who failed to condition were aged ≤2.5 yr. Only one of the two hearing aid users who failed to condition was under 2.5 yr of age at testing. The 4-yr-old child using an implant, who failed to condition, is reported to respond very inconsistently during the mapping of electrode thresholds. A neurological assessment has identified a disorder consistent with leukodystrophy, but that is not neurodegenerative.

There were also children in the present procedure who could be conditioned but could not complete the task over three 20 minute sessions. Audiological reports indicated that these children either were unable to condition to conventional play audiometry or were unable to perform consistently in play audiometry. One of the children was receiving medication for an attention deficit disorder. The normally hearing and hearing-impaired children who failed to complete the task generally exhibited inattentiveness, poor waiting behavior, and frequent loss of conditioning.

The attrition rates for the SFT are very encouraging, especially given that the task is cognitively more demanding than play audiometry. There is transfer of learning required in the play audiometry task in that acuity is tested at different frequencies. This transfer is simple because the new stimulus remains a pure tone and quickly can be made loud enough for the child to hear. In the SFT task the child needs to be able to transfer the conditioned response taught during training to each new change in speech stimuli.

Reasonably high attrition rates due either to inability to condition or inability to complete the task also have been reported for infant testing with the VRISD task and for assessment of primary-aged children with the change/no change task. Nozza et al. (1991), in a study of 7- to 11-mo-old infants’ speech sound discrimination-in-noise thresholds, found that 8/34 infants could not be conditioned and two could not complete testing. Carney et al. (1993) reported that, of 48 profoundly hearing-impaired children assessed, four children (mean age of 5.2 yr) failed to condition to the change/no change task and 12 children (mean age of 7.4 yr) could not complete the 15 trials for all nine contrasts tested. Similarly, Osberger, Robbins, Miyamoto, Berry, Myres, Kessler, and Pope (1991) found that some hearing-impaired children could not complete the 15 trials for all nine contrasts assessed in the change/no change task.

Ease and Efficiency of Testing

On average, hearing-impaired children took a greater number of sessions than normal hearers to complete the task. This is to be expected given that they often failed some contrasts and needed to be given at least one trial of a previously passed contrast after any failed contrast to provide positive reinforcement and to ensure that they were still conditioned.

Lack of transfer was found to be present for a small number of normally hearing and hearing-impaired children, including some 4-yr-old hearing-impaired children. The decision to look for transfer difficulty during the testing of the first four contrasts was supported by the finding that, for each of the six normally hearing children who had difficulty in generalization, the difficulty occurred before the presentation of the fifth contrast. It may have been anticipated that the presence of a hearing loss, which decreases the saliency of the speech contrasts, could compound the problem of generalization across different speech contrasts. This did not appear to be the case, and, furthermore, the degree of hearing loss did not appear to affect the likelihood of transfer difficulty for the children in this study. Several children using hearing aids who experienced difficulty in generalization had PTAs of 72 dB HL or less. The fact that the contrasts that were predicted to be easiest for profoundly hearing-impaired children were presented first probably prevented the degree of hearing loss from being influential.

Future research could investigate the usefulness of conditioning the child on the task with two or more contrasts in the training phase rather than with just one highly salient speech contrast. Alternatively, it may be worthwhile to provide practice trials on each new contrast as did Carney et al. (1993).

It was apparent during the field trials with the normally hearing children and the hearing-impaired children that many 3- and 4-yr-old children could be assessed without an assistant tester present. In contrast, an assistant usually was essential with the 2-yr-old group to maintain the child’s attention and motivation and to facilitate game changes. As anticipated, some 2- and 3-yr-old children were tentative in performing with the required motor response but did give reliable behavioral responses. The latter responses were recorded for only one normally hear-
Speech Discrimination Performance

Normally Hearing Children • All contrasts were discriminated successfully by each of the 24 normally hearing children who were able to complete the task, and older children did not perform better than younger children. These findings are relevant only for the particular set of speech contrasts used in this study. Although findings of previous investigators have suggested developmental changes in auditory discrimination of speech sounds in normally hearing children (Elliott, Busse, Partridge, Rupert, & DeGraaff, 1986; Elliott, Longinotti, Meyer, Raz, & Zacker, 1981; Sussman & Carney, 1989), this study’s finding is not discrepant with these. Unlike the previous findings, this study’s results were based on a small number of six presentations per contrast and did not focus on the measurement of fine discriminations between synthetic CV syllables presented along a continuum.

Hearing Aid Users • Contrast type rather than age influenced the discrimination performance of those children using hearing aids who completed the task. As found by Boothroyd (1984), for severe and profoundly deaf children tested on the IMSPAC, average performance for hearing aid users in the present study was superior for the easier speech contrasts: the simple intonation contrast, the vowel duration contrast, and the contrast varying both vowel height and frontedness. Osberger, Robbins, Miyamoto, Berry, Myres, Kessler, and Pope (1991) found that children using hearing aids who had useful high-frequency residual hearing and children using the 22-electrode cochlear implant obtained scores >80% for these easier contrasts as well as for vowel place and consonant manner contrasts. The latter contrasts were probably easier to discriminate than the vowel place and manner contrasts in the present study. Unlike the /ta/sa/ contrast in the present study, there was an overall syllable duration difference of 93 msec for the /da/za/ contrast in the Osberger study, which could have been used as a secondary cue to detect the phonetic change. The American /b/ vs. /v/ vowel place contrast also involved a first formant difference of 215 Hz, whereas the first formant difference in the /b/ vs. /v/ contrast in the present study was negligible (Carney et al., 1993).

There was, as expected, a significant effect of hearing loss, with profoundly hearing-impaired children performing more poorly overall than severely hearing-impaired children, and severely hearing-impaired children, in turn, performing worse than the children with a mild to moderate hearing loss. This effect was evident only for the discrimination of harder speech contrasts that required the use of finer spectral information and/or fine temporal processing. The pattern of results for the different contrasts for children with varying degrees of hearing loss was in reasonable agreement with the known acoustic properties of the speech contrasts. It must be remembered that although length differences between /m/ and /h/ and between /s/ and /t/ were preserved, there were minimal overall syllable length differences for these contrasts. The child would need to be able to discriminate the frequency differences between these contrasts to use the length differences on the initial phonemes.

A regression analysis also indicated that hearing loss (PTA) accounted for a significant amount of the variance in individual scores, especially for the speech contrasts requiring use of finer spectral cues. No other variables, such as aid experience, age at aid fitting, and age at time of testing, accounted for a significant portion of the variance. There were too few children in the total communication setting compared with an oral/aural setting to determine the effect of mode of communication on discrimination performance. The PTA slope coefficients for performance on the contrasts /ba/ma/, /ba/ga/, /ba/pa/, /bi/ bu/, and /ta/sa/ suggested that for every 10 dB of hearing loss, discrimination would drop by approximately half a mark out of a possible total of six marks. The PTA slope coefficient for the percentage of contrasts successfully discriminated indicated that for a change in PTA of 20 dB the percentage of contrasts passed would change by about 14% or approximately one contrast.

Cochlear Implant Users • The average performance of the 10 cochlear implant users was similar to that of the hearing aid users for the three easier contrasts: /ba/ba/, /ba/ ba/., and /ba/bi/. For the other contrasts, average performance was approximately midway between the performance of the severely deaf hearing aid users and the performance of the profoundly deaf hearing aid users. The group of implant users in this study, however, is only small, and these findings cannot be generalized to a wider population of implant users. Factors such as duration of implant experience and duration of deafness did not account for variability in the performance of this group of implant users. This finding may be explained partly by the limited variation in either of these factors for this group. It is of note that psychological assessments indicated cognitive deficits for two of the four implanted children who did poorly (failed between five and seven contrasts). The
third child has been reported to display inconsistent device use. An electrical auditory brain stem response with the implant for the fourth child who did poorly indicated abnormal retrocochlear functioning and suggested deficiency of nervous tissue in the cochlea. This child uses the wide mode of stimulation (bipolar + 4). Four of the remaining six implant users passed all of the eight contrasts, performing similarly to the mild to moderate hearing aid users.

Validity

The test results are intended to reflect the child’s ability to perceive phonetic changes. A child may, however, detect that the syllable has changed on the basis of duration, F0, or perceived loudness. The overall length of stimuli within a contrast did not differ by more than 18 msec. Studies have revealed that the duration difference limens for acoustic stimuli decrease from approximately 35% at 10 msec to 14% at 600 msec for normally hearing adults (Abel, 1972; Creelman, 1962). It was, therefore, most unlikely that a child could detect a change in a syllable on the basis of overall duration. Syllables within a contrast did not differ in terms of F0 (start or end) by more than 7 Hz. Olsho, Koch, and Halpin (1987) report frequency difference limen of 3 to 4% in normally hearing infants. In light of this finding, it is possible but unlikely that children in this study could use F0 as a cue to the presence of a phonetic change. There is also the possibility that a child who fails a contrast (scores <4/6) actually may perceive the changes in acoustic cues but fail to respond because the contrast lacks salience or immediate significance for him or her. For example, the contrast may lack phonological significance if the phonemes in the syllable occur infrequently in the speech production of the child. This issue of salience could be addressed in longitudinal developmental studies.

Subjectivity

Another issue to consider is whether the child’s performance is an accurate indication of his or her speech discrimination ability. All testing with very young children is characterized by subjectivity. The skills of the tester and the rapport achieved between tester and child can be crucial in determining whether or not a child conditions to the task, sustains attention, or performs at maximum capacity. Aspects of the current procedure, such as the tester receiving masking noise and the computer control over timing of stimulus presentation, were incorporated to reduce subjectivity in administration and scoring (see Method section). The order of presentation of stimuli was controlled in this task to encourage the child to perform optimally.

Test-Retest Reliability

The procedure has high reliability and good stability. The latter indicates that children not only are ranked similarly on a retest but score similarly to the initial test, with an average difference (retest score minus test score) of one-fifth of a mark or 3.5%. This compares favorably with the 4.5% mean difference score for the NU-CHIPS, a four-alternative word recognition task for hearing-impaired children (Elliott & Katz, 1980).

False Positive Rates

It was promising that the false positive rates for the three groups of children were very low. On three occasions the probability that a false positive response would occur in the hit window exceeded 0.26, but in two cases the contrasts were discriminated successfully with scores ≥5/6, which still allowed a confidence level of 0.05. In the third instance the contrast was not discriminated successfully. It should be remembered that the confidence levels for successful discrimination for most children were much better than the 95% level. The average probability of a false positive occurring in the hit window for each of the three groups of children was less than 0.048, and the confidence level for scores ≥4/6 with this low probability is 99.9%.

Future Uses of the Test

This procedure has the potential to be both a useful clinical and research tool. It could be used in combination with other assessments to gauge the auditory discrimination capacity of the child during a preimplantation and a postimplantation period. Test information obtained over time on a sufficiently large number of implanted children ultimately could be used to assist in determining candidacy for implantation. Test information also could help in planning habilitation programs.

The technique could be used to measure speech feature discrimination longitudinally in severely/profoundly deaf children over time. Many studies have looked at the development of speech feature discrimination in infants with normal hearing (Eilers & Minifie, 1975; Eilers et al., 1977) and at the development of auditory perceptual skills in children around 6 to 10 yr of age (Elliott et al., 1986; Sussman & Carney, 1989). In contrast, there has been a scarcity of research on the development of
speech feature discrimination in preschool severe-profoundly hearing-impaired children. In fact, the auditory skills of many hearing-impaired children in this age group cannot be determined reliably given that conventional assessments require knowledge of particular vocabulary or prerequisite speech production ability. The stimuli in the SFT can be chosen and instrumentally controlled for specific testing purposes. For example, it may be of interest to determine which acoustic cue is being used predominantly by the young implanted child in discriminating place of articulation contrasts or voicing contrasts. It is important to note that many of the severely hearing-impaired children and implant users passed all contrasts, indicating the need to include more difficult contrasts requiring the use of fine spectral cues in a future test battery.

Future work could assess the validity of the procedure by investigating the correlation between overall performance on this task with performance on word recognition tasks and even sentence tasks with older children.

CONCLUSIONS

This study has addressed the need for a speech perception procedure for severe-profoundly deaf 2- to 4-yr-old children that which is independent of language and speech production ability. The new SFT is a modification of existing procedures used and appeared to be a reasonably efficient and reliable test for the majority of 3- and 4-yr-old severely-profoundly deaf children tested and for about 50% of 2-yr-old children tested. Performance on the task was influenced by PTA hearing loss but not by age at testing for the particular speech contrasts tested. The procedure has the potential to be used as a clinical and research tool in the assessment of very young, severely-profoundly deaf children using hearing aids or cochlear implants.

ACKNOWLEDGMENTS:

This work was supported by the Lions Club International, the Australian Bionic Ear and Hearing Research Institute, and the Cooperative Research Centre for Cochlear Implant, Speech, and Hearing Research. Many thanks to the parents of the children involved and the teachers at Taralga Early Intervention Program, Furlong Park School and Pre-School for Deaf Children, and Princess Elizabeth Junior School. We would like to especially thank Dr. Peter Blamey, Professor Barbara Cone-Wesson, and Dr. Colette McKay for their input on research design. We are also very grateful to John Seymour for writing the computer program for the test. Finally we would like to thank Alison King for her role in recording and editing the speech stimuli.

Address for correspondence: Ms. P. Dawson, Australian Bionic Ear and Hearing Research Institute, 384–388 Albert St., East Melbourne 3002, Australia.

Received July 7, 1997; accepted May 21, 1998.

REFERENCES


**REFERENCE NOTES**

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Dawson, P. W.; Nott, P. E.; Clark, Graeme M.; Cowan, Robert S. C.

Title:
A modification of play audiometry to assess speech discrimination ability in severe-profoundly deaf 2- to 4-year-old children

Date:
1998

Citation:

Persistent Link:
http://hdl.handle.net/11343/27514