SPEECH DISCRIMINATION IN PRESCHOOL CHILDREN: A COMPARISON OF TWO TASKS

SUSAN MENARY
McGill University

SANDRA E. TREHUB
University of Toronto

JAMES McNUTT
McGill University

Four-year-old children were tested for their discrimination of the following word pairs: rope/robe, seasteed, pick/pig, ice/eyes, and house/mouth. Two discrimination tasks were used: a picture-pointing task, which typifies discrimination tasks currently in use with young children, and an operant technique which exemplifies procedures used in the testing of infants. Both discrimination tasks yielded roughly comparable levels of performance. All word pairs were found to be discriminable, but performance on set seed and house/mouth was inferior to that of the other word pairs. In a limited sampling of production data, misarticulations were observed in the final consonants of ice, eyes, mouse, and mouth.

The speech perception skills of infants and young children have been a topic of growing research interest in the last decade. Despite the volume of published research, our understanding of the speech perception process in the first few years of life remains limited. Researchers have found infants to be capable of perceiving many of the consonant and vowel contrasts that are phonemically relevant in English (e.g., Eilers, 1977; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Moffitt, 1971; Trehub, 1973; Trehub & Babinovitch, 1972). In contrast, the speech perception abilities of 1- to 3-year-old children have been found to be incomplete (Eilers & Oller, 1976; Garnica, 1973; Locke, 1971; Shvachkin, 1973) and to improve gradually until about 9 years of age.

© 1982, American Speech-Language-Hearing Association
The apparent differences between infants and young children have been attributed to differences in methodology and task demands which characterize studies of these different age groups (Barton, 1978; Trehub, 1979; Trehub, Bull, & Schneider, 1981), raising the possibility that quite different skills are being observed. On the surface, at least, it would appear that the developmental course of speech perception could be revealed more appropriately through the application of uniform techniques across a broad age range.

The two most commonly applied infant research techniques, the high-amplitude sucking procedure (Elmas et al., 1971; Morse, 1972; Trehub, 1973) and the heart rate habituation-dishabituation procedure (Miller & Morse, 1976; Miller, Morse, & Dorman, 1977; Moffitt, 1971), have not been used to explore speech perception skills in the preschool years. In the case of the high-amplitude sucking technique, the response measure is obviously inappropriate beyond early infancy. The heart rate measure, although potentially applicable beyond the infant period, has been found at times to be insensitive as a discrimination tool, even for infants (Leavitt, Brown, Morse, & Graham, 1975). Recently, however, Elbers, Wilson, and Moore (1977, 1979) have pioneered an operant head-turning procedure (termed VRIISD or Visually Reinforced Infant Speech Discrimination) which has been used successfully with 6- to 14-month-old infants and appears potentially applicable during the preschool years.

Although the infant speech discrimination techniques require an infant to indicate, in some way, that a change in stimulation has occurred, the techniques typically in use with children 1-3 years of age embody considerably different demands, such as the learning of nonsense names for objects (e.g., Elbers & Oller, 1976; Garnica, 1973; Shvachkin, 1973). For children beyond 2 or 3 years of age the assessment of speech discrimination frequently involves same/different judgments (e.g., Templin, 1957; Wepman, 1958) which have been shown to generate confusion in young children (Beving & Elben, 1973; Donaldson & Wales, 1970). In addition, picture-pointing tasks (e.g., Barton, 1976; Goldman, Fris-toe, & Woodcock, 1971; Haggard, Corrigall, & Legg, 1971; Higgs & Hodson, 1978; Locke, 1971; Menyuk & Anderson, 1969) with known and/or taught words are often used.

The principal purpose of the present investigation was to compare the discrimination performance of 4-year-old children on infant and child discrimination tasks. The VRIISD technique was selected to represent infant tasks because of its potential ease of application beyond the infant period. A picture-pointing task served as the child discrimination technique and as a replication of Higgs and Hodson. The stimulus pairs were CVC or VC words, with members of each pair differing in their final consonants. Final consonant differences were selected because they appear to be more difficult for children (e.g., Atchison & Canter, 1979; Shvachkin, 1973), and they have been studied relatively infrequently. These final consonants differed in voicing or place of articulation. Voicing differences were thought to be particularly relevant because of reports of infants' discrimination of voicing contrasts (Elmas et al., 1971; Trehub & Babinovitch, 1972) and young children's difficulties with such contrasts (Garnica, 1973; Greenlee, 1978; Shvachkin, 1973). Moreover, comparison data from 4-year-old children on final voicing contrasts in a picture-pointing task were available (Higgs & Hodson, 1978). Frivolities differing in place of articulation were included because of infants' and young children's perceptual difficulties (Elbers & Oller, 1976; Elbers et al., 1977) and children's articulation errors (Elbers & Oller, 1976) with these sounds.

**METHOD**

**Subjects**

The subjects were 11 4-year-old children selected by telephone solicitation from a suburban area adjacent to Toronto. There were 5 girls and 6 boys ranging from 4 years to 4 years, 5 months of age with a mean age of 4 years, 1 month. The children had no obvious speech or language difficulty or other developmental delay and no observable cold or illness on the test day.

**Apparatus**

The experiment was carried out in an IAC double-walled booth. The booth contained two chairs on opposite sides of a table where the experimenter and child sat facing each other during testing. A Nova-6 (Radio Shack) loudspeaker and a box faced with dark Plexiglas were located at a 45° angle to the child's left at eye level. A 3-dimensional toy (a Pluto dog or toy train) was located inside the box. When electronically activated, the box was backlit and the toys moved and sounded.

The equipment outside the booth was operated by a second experimenter, and consisted of a Tandberg 9200XD tape recorder, a Marantz model 1070 amplifier, and a specially designed electronic unit that recorded responses and determined the trial type (experimental or control).

**Stimuli**

The auditory stimuli were five single-syllable CVC and VC word pairs that differed in their final consonants. The words were: rope/rope, sheet/heed, pick/pig, ice/evers, and mouse/mouth. Members of the first four word pairs, which had been used by Higgs and Hodson (1978) with 4-year-olds, differed in final consonant voicing, and members of the fifth differed in final consonant place of articulation. Examples of the five word pairs were spoken repeatedly by a native female English speaker and were recorded on a Tandberg tape recorder. The stimuli were then digitized by computer, and word pairs with
similar overall duration and similar frequency contours were chosen. The duration values for the word pairs can be seen in Table 1. Note that the voicing cue in the first four word pairs was associated with a difference in the duration of the preceding vowel, as was noted in previous research (Denes, 1955; Raphael, 1972). Rather than altering these cues, we decided to keep overall duration differences at a minimum. None of the differences between members of a pair exceeded 100 msec.

Two tapes were made using these word pairs. For the first tape, which was used in VRISD testing, each word pair was placed with one word on channel 1 (e.g., rope) and its contrasting member (e.g., robe) on channel 2, with simultaneous onset. The words were repeated on this tape 400 times at a rate of one per sec; this was done for each of the five word pairs. The tape for the picture-pointing task was a single randomized order of the 10 words, with each word being repeated five times in succession (i.e., one after the other). The rate of presentation was one word per 10 sec.

Colored drawings representing each of the 10 words were made and each drawing was affixed to a 20.3 x 30.5 cm. cardboard sheet. Rope was pictured by a coll of rope, robe by a bathrobe, seat by a car-seat, seed by brown seeds, pick by toothpicks, pig by a pink pig, ice by ice cubes, eyes by a pair of eyes, mouse by a brown field mouse, and mouth by an open human mouth.

Procedure

Testing was divided into two phases—VRISD and picture-pointing. Half of the time VRISD was tested first and the other half, picture-pointing was first.

VRISD testing. The VRISD paradigm, as reported in Eilers, Wilson, and Moore (1977, 1979) involves conditioning a child to respond to a change in a repeating background stimulus (in this case, a word) to receive visual reinforcement (in this case, seeing an animated toy for a short period). From the child’s point of view, a game was played, the object of which was to turn on the toy inside the box, a feat that could be accomplished by pointing to the box when the background sound changed. A training session using the nonminimal contrast cup and ball preceded the test session. The experimenter began by modeling the desired response on two successive trials. The child was then required to achieve a training criterion of three successive correct responses. Typically, the criterion was attained with error-free responding. In only two cases were more than three trials required; in one instance this resulted in five training trials, in the other six.

During testing, the child was exposed to one member of a word pair (e.g., rope) presented repeatedly at 65 dB SPL (A) (measured by a Bruel & Kjaer sound level meter) until the child was judged by the experimenter (E1) to be quiet and attentive. At this point, E1 pressed a button to signal the second experimenter (E2) to begin a trial. An experimental trial consisted of four repetitions of the contrasting member of the word pair (e.g., robe) followed by a return to the original repeating word. The child was reminded occasionally during intertrial intervals to point when the sound changed. If the child responded (i.e., pointed) during the word-change interval, the toy was activated for 5 sec. Responding between trials was ignored. Control or no-change trials were included to assess the possibility of random responding, and consisted of a monitored 4-repetition interval of the background word. There were 5 randomly presented change or experimental trials and 5 no-change or control trials. This procedure was repeated for each of the five word pairs. The order of presentation of word pairs and experimental and control trials was randomized and known only to E2, who operated the equipment outside the booth. E2 inside the booth wore headphones with music that masked all other sound inside the booth. E2 operated a control unit for communicating with E2 about trial initiation and responses and was blind to the occurrence of change or no-change trials.

The data consisted of responses on 10 trials per word pair, 5 change and 5 no-change.

Picture-testing. The physical set-up was identical to that for VRISD testing, except that E2 controlled the presentation of sounds within the booth with a remote control unit and E2 did not wear headphones. The child was first required to name all pictures correctly, and unknown words were taught. The child was then instructed to point to the picture signalled by the recorded voice. There was a brief training period with the words cup and ball to check on the child’s comprehension of instructions. All children responded immediately. Pictures were presented in pairs, with side of presentation of the correct picture randomized. A trial began with a single

<table>
<thead>
<tr>
<th>Component</th>
<th>rope</th>
<th>robe</th>
<th>seat</th>
<th>seed</th>
<th>pick</th>
<th>pig</th>
<th>ice</th>
<th>eyes</th>
<th>mouse</th>
<th>mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial c</td>
<td>89.7</td>
<td>79.8</td>
<td>103.6</td>
<td>107.7</td>
<td>48.0</td>
<td>83.6</td>
<td>18.1</td>
<td>28.1</td>
<td>73.9</td>
<td>56.4</td>
</tr>
<tr>
<td>Vowel</td>
<td>261.9</td>
<td>365.3</td>
<td>185.3</td>
<td>224.3</td>
<td>129.8</td>
<td>385.4</td>
<td>313.1</td>
<td>451.9</td>
<td>308.0</td>
<td>331.1</td>
</tr>
<tr>
<td>Silence</td>
<td>111.0</td>
<td>64.4</td>
<td>49.5</td>
<td>63.4</td>
<td>36.0</td>
<td>24.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>165.5</td>
</tr>
<tr>
<td>Final c</td>
<td>45.2</td>
<td>56.4</td>
<td>98.4</td>
<td>64.5</td>
<td>190.7</td>
<td>47.7</td>
<td>264.0</td>
<td>208.0</td>
<td>226.0</td>
<td>151.9</td>
</tr>
<tr>
<td>Total</td>
<td>510.8</td>
<td>566.3</td>
<td>525.8</td>
<td>549.3</td>
<td>402.3</td>
<td>410.9</td>
<td>593.2</td>
<td>688.7</td>
<td>609.0</td>
<td>690.9</td>
</tr>
<tr>
<td>Overall difference</td>
<td>55.5</td>
<td>24.0</td>
<td></td>
<td></td>
<td>8.6</td>
<td>93.5</td>
<td>81.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vowel difference</td>
<td>103.4</td>
<td>38.8</td>
<td></td>
<td></td>
<td>125.6</td>
<td>138.8</td>
<td>23.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final c</td>
<td>8.2</td>
<td>33.9</td>
<td></td>
<td></td>
<td>143.0</td>
<td>56.0</td>
<td>74.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Duration (msec) of words and word components.
presentation of the target word and terminated when the child pointed to a picture. There was one prerandomized order of presentation for all subjects in the picture-testing session, and each child was presented with each word 5 times, giving a total of 50 trials (10 per word pair). During the initial picture-naming phase the child’s productions (typically one per picture) were transcribed. A production error was scored only when the target sound, the final consonant, was misarticulated.

RESULTS

For the VRISD task, the proportion of pointing responses occurring during experimental (change) and control (no change) intervals was compared for each of the five word pairs. Z tests revealed highly significant differences ($p < .001$) for all word pairs, indicating that the contrasting members of each word pair were discriminated by the 4-year-olds as a group. Correct and incorrect responses on each word pair are presented in Table 2. It is apparent that performance on the seat/seed and mouse/mouth pairs is considerably worse than on the remaining pairs.

Table 2. Responses on VRISD task.

<table>
<thead>
<tr>
<th>Word pair</th>
<th>Number correct</th>
<th>Number incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>rope/robe</td>
<td>106</td>
<td>4</td>
</tr>
<tr>
<td>seat/seed</td>
<td>78</td>
<td>32</td>
</tr>
<tr>
<td>pick/pig</td>
<td>107</td>
<td>3</td>
</tr>
<tr>
<td>ice/eyes</td>
<td>103</td>
<td>7</td>
</tr>
<tr>
<td>mouse/mouth</td>
<td>79</td>
<td>31</td>
</tr>
</tbody>
</table>

Correct and incorrect responses on the picture-pointing task are illustrated in Table 3. Again, it can be seen that performance on the seat/seed and mouse/mouth pairs is inferior to the other contrasts. Performance on the separate words of each pair is shown in Figure 1. It appears that children more readily identified seat than seed and were also superior on pig compared to pick and on mouth compared to mouse.

Table 3. Responses on picture-pointing task.

<table>
<thead>
<tr>
<th>Word pair</th>
<th>Number correct</th>
<th>Number incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>rope/robe</td>
<td>94</td>
<td>16</td>
</tr>
<tr>
<td>seat/seed</td>
<td>84</td>
<td>26</td>
</tr>
<tr>
<td>pick/pig</td>
<td>97</td>
<td>13</td>
</tr>
<tr>
<td>ice/eyes</td>
<td>99</td>
<td>11</td>
</tr>
<tr>
<td>mouse/mouth</td>
<td>82</td>
<td>28</td>
</tr>
</tbody>
</table>

Data from the VRISD and picture-pointing tasks are shown in Figure 2. For each task the maximum score is 10. In the case of the VRISD task, a score of 10 would be achieved by having 5 correct pointing responses on change intervals and 5 correct instances of no pointing on no-change intervals. For the picture-pointing task a score of 10 would be obtained when a subject pointed to the correct picture on the 5 trials of each of the members of a word pair. It can be seen that overall performance is roughly comparable on the two tasks in the sense that the mean scores on both tasks never differ by more than one point. An analysis of variance (task × word pair) revealed a highly significant effect of word pair, $F(4, 40) = 8.839, p < .001$, but no effect of task ($F < 1$) and no task × word pair interaction, $F(4, 40) = 1.364, p > .20$. Newman-Keuls tests confirmed that mouse/mouth and seat/seed did not differ from each other but differed significantly from the remaining pairs ($p < .01$) which did not differ from each other.

The VRISD data for individual subjects were examined with Fisher’s tests. Of the 11 children in the sample, 10 performed significantly better than chance on seat/seed and pick/pig, and 9 performed above chance levels on ice/eyes. In contrast, only one child performed beyond chance levels for rope/robe and three for mouse/mouth. The picture-pointing data for individual children were analyzed with binomial tests. The num-
bers achieving performance levels significantly exceeding chance were 11 for pig; 9 for rope, robe, sed, ice, and eyes; 8 for mouth; 7 for pick; 5 for mouse; and 4 for seat.

Production data comprised a single sample of each word from each child and, as noted, an error was scored only when the target phoneme—the final consonant—was misarticulated. According to this criterion, all subjects had correct productions of rope, robe, sead, and pick/zip. In contrast, ice and eyes were incorrectly articulated by 4 of the 11 children, mouth by 3 children, and mouse by 5 children. Two common types of substitutions were seen: (a) /b/ for /s/ or a frontal lisp in ice and mouse (and its cognate /s/ for /z/ in eyes), and (b) /l/ for /w/ in mouth. In only two instances was production nondistinctive between members of a word pair. One child produced an identical utterance (/maouth/) for mouse and mouth, and another child did so (/aio/) for ice and eyes.

**DISCUSSION**

Results of the present investigation reveal that the VRISD technique is applicable to 4-year-old children in addition to the infant populations with whom it has been used previously (Eilers, Gavin, & Wilson, 1979; Eilers et al., 1977, 1979). Moreover, the VRISD technique was found to yield comparable performance, based on group data, to the more traditional picture-pointing task. Thus, although a cursory task analysis might suggest that the VRISD technique poses fewer cognitive demands than the picture-pointing task, the present performance data did not reflect this. It is possible that the picture-pointing task was, along some dimensions, more difficult than VRISD for these children but that this difficulty was balanced by greater familiarity with everyday tasks involving pointing to named pictures. Alternatively, the high level of performance on all words may have masked differences in the relative difficulty of the tasks. One could resolve the latter issue by selecting more difficult discriminations or by studying a younger age group, for example, 2- or 3-year-old children.

The performance of individual children revealed some intertask variability, with particular children performing more readily on the VRISD task and others showing superior performance on the picture-pointing task. This underscores the utility of alternative assessment techniques in clinical applications. For such applications the two procedures, as used in the present investigation, yield individual performance data that are statistically analyzable. Where individual assessments are of principal interest, however, it would be preferable to increase the number of observations on each word pair, decreasing, if necessary, the total number of word pairs.

Based on Higgs and Hodson's (1978) data, performance on the perceptual tasks, at least the voicing pairs, was expected to be inferior to that actually obtained. The substantially greater error rate of Higgs and Hodson may be attributable to distraction associated with their simultaneous presentation of eight pictures (relevant to the four word-pairs used) in contrast to the present practice of presenting two pictures on each trial. On the other hand, it is unlikely that their use of the carrier phrase "Where is the x?" could account for observed discrepancies.

In the present study rope, robe, pick/zip, and ice/eyes were found to be relatively easy discriminations whereas seat/seed and mouse/mouth posed some difficulties for children on both discrimination tasks. In the case of seat/seed, perceptual difficulties may be related to the small durational difference between the vowel preceding the final consonant relative to the other voicing pairs (see Table 1). This interpretation is consistent with the view of vowel duration as the strongest cue for final consonant voicing (Abbs & Minifie, 1969; Greenlee, 1978; Raphael, 1972). In fact, it has been suggested that vowel duration is the child's first productive marker to signal final consonant voicing (Naezer, 1970; Velten, 1943). In contrast, similarities in overall duration, as in pick/zip, or final consonant duration, as in rope/robe, did not appear to affect discriminability. In the case of mouse/mouth, perceptual difficulties may be related to the high-frequency, low-intensity differences that cue this contrast and/or the absence of vowel-duration cues. Note that for ice/eyes these high-frequency, low-intensity fricative cues are supplemented by vowel duration cues.

The limited production data revealed errors that failed, for the most part, to mirror perceptual problems. Productions were largely distinctive between members of the word pairs, suggesting that the articulation difficulties seen here were not linked to perceptual confusion. Even for studies with a more adequate sampling of production, the relation between perception and production has remained elusive, both for normal children (Edwards, 1974; Eilers & Oller, 1976; Haggard et al., 1971) and for children with articulation or language problems (Rees, 1973; Weiner, 1967).

Finally, although the present study clearly demonstrates the feasibility of extending the VRISD technique to 4-year-old children, it remains to be determined whether this procedure can be used effectively with children in the intervening period between infancy and 4 years. If VRISD is applicable through that age span, it would promote the systematic investigation of numerous unresolved issues concerning the development of speech perception skills.

**ACKNOWLEDGMENTS**

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to S. E. Trehub. The authors are grateful to D. Douchin and B. Schneider for their statistical advice and to L. Thorpe and M. Endman for assistance in testing children.

**REFERENCES**


Received August 5, 1980
Accepted February 17, 1981

Requests for reprints should be sent to Sandra E. Trehub, Centre for Research in Human Development, Erindale College, University of Toronto, Mississauga, Ontario, Canada L5L 1C6.