Discrimination of Formant Transitions by Down Syndrome and Normally-Developing Infants

Rebecca E. Eilers, Debra A. Moroff, and Roberta H. Turner

INTRODUCTION

Language Problems of Down Syndrome Children

Down Syndrome is one of the most prevalent causes of severe retardation. Individuals with Down Syndrome (DS) show general developmental lags, but their retardation in speech communication seems to be more severe than would be predicted by mental age alone (Akin & Aslin, 1969).

The communication problems of DS children are most evident in the auditory-vocal channel. Belmont (1971) concluded that in DS the auditory channel is weaker than the visual and the auditory-vocal channel is particularly deficient. Performance by DS children was found by Schellenberg (1968) to be poorer on experimental tasks that tapped the auditory-vocal channel than on tasks that tapped the visual-vocal, auditory-motor, or visual-motor channels. Other evidence indicates that gesture-communication is superior to vocal communication in DS children (Biolok & Share, 1966) and that reception of gestural symbols by DS children may be better than the reception of spoken words (Oiler, Coleman, & Eilers, 1978). Finally, Schlinger & Gottsleben (1957) and Smith & Stoel-Gammon (1983) have reported that DS children have substantial difficulties in speech articulation.

In spite of the observations that DS children have difficulties in acquiring spoken language, little is known about the nature of the deficits that result in the relative language delay. One possible contributing factor - conduction hearing loss sometimes complicated by sensory-nervous damage, is well documented in DS children and infants. Several investigators have restricted a high incidence of middle-ear pathology (60%) in samples of asymptomatic DS children (Schwartz & Schwartz, 1978; Bal­lany et al., 1979). Greenberg et al. (1978) reported depressed speech reception thresholds in DS populations, probably a consequence of the high incidence of loss. In addition, Walth, Salvisber, & Auf der Maur (1973) have suggested that auditory memory deficits might play a role in language-learning difficulties of DS children.

Auditory-Speech Processing in DS Children

In addition to peripheral sensory problems, central auditory processing may also be deficient in DS children. In a recent study (Eilers & Oller, 1980) of a retarded group of children (including DS), discrimination of a steady-state speech stimulus (vowel contrast) was comparatively better than discrimination of dynamically-changing stimuli (consonantal contrast in a consonant-vowel syllable). Interestingly, Tallal (1975) and Tallal & Piercy (1974, 1975) demonstrated that young, non-retarded children with specific language problems may have similar central deficits. These language impaired children have more difficulty discriminating rapid formant transitions (spectral changes accompanying articulatory movement) than slowly varying or steady formants. In these two language-delayed populations, children’s discriminative performance on rapidly changing spectral speech cues was relatively poor compared to their performance on slowly changing or steady-state speech cues. A deficit in rapid processing might result in severe language problems associated with deciphering the speech signal, especially in rapid conversational speech. Rapidly changing spectral information embedded in formant transitions accounts for a wide variety of speech contrasts, including place of articulation of consonants, distinctions of stops, glides and vowel sequences, and vowel and semivowel quality. In rapid speech, steady-state vowel resonance seldom occurs. Thus far, only one study has addressed the hypothesis that some specific language difficulties in DS may be explained by a deficit in rapid-spectral processing.

In a preliminary attempt to evaluate whether DS infants have difficulties in discriminating speech information contained in formant transitions, DS and normally-developing infants were presented with three pairs of speech contrasts in a discrimination study (Eilers et al., 1985). The first speech contrast pair, /ba/ vs. /ga/, contained at least 25 ms transitions between the consonant and vowel elements. The members of the pair...
differed only in the direction of the second formant during the 25 ms transition. The second pair, /wa/ vs. /ga/, contained normal duration transitions (75 ms) for semi-vowel to vowel syllable transition. The third pair, /ba/ vs. /ga/, contained atypically slow transitions (225 ms) yielding the percept of two distinct vowel nuclei. Although stimulus pairs differed in a second formant transition duration, spectral information was identical in all three pairs. Results indicated that there was no overall significant difference between DS and normally-developing babies' discrimination of the stimulus pairs. DS infants performed significantly better than the slowest transition (225 ms) pair than they did on either the very rapid (25 ms) pair or on the standard (75 ms) pair. In contrast, normally-developing infants performed equally well on the 75 ms and the 225 ms pairs which were both discriminated better than the 25 ms pair. Thus, DS infant performance was poorer for rapid transition durations typical of speech-like syllables than for slowly presented spectral information. The present study was designed as a follow-up to explore further the ability of DS infants and young children to discriminate rapid spectral speech events.

METHOD

Subjects

Nine full-term normally-developing infants (28-31 weeks of age) and nine DS infants were selected for the study. The DS infants were selected from the High Risk Follow-up Clinic at the Mailman Center for Child Development. Several of the DS infants were tested with the VRISD discrimination procedure (Thompson, Wilson, & Moore, 1979) for several months before they demonstrated appropriate orientation to the auditory stimuli, a prerequisite for task success. The DS infants were entered in the study as soon as a parent could schedule an appointment and at the point at which normal orientation by manipulating a set of quiet toys. In all cases the adult holding the infant and the experimenter listened to masking music over headphones throughout the testing session. Four adults who participated in discrimination and identification testing of the stimulus pairs had no difficulty correctly identifying each of the six stimuli. Discrimination scores were 100% for all adult perceivers.

Stimuli

Two types of stimuli were used: training and test. The training stimulus contrast pair consisted of five tokens for each of the syllables /ba/ and /ga/ produced by a male phonetician and matched pairwise for overall duration, overall amplitude, peak amplitude, mean fundamental frequency, and peak fundamental frequency. These stimuli differed in the formant structure of the vowel and had been used successfully in the past as a VRISD training contrast with DS infants (Eilers et al., 1985).

The experimental test stimuli consisted of three pairs of syllables, synthesized using the Klatt (1980) routines. The stimuli of the first syllable pair (also used in the previous study), /ba/ and /ga/, differed from each other only with respect to the slope of the second formant transition. In natural speech articulatory terms, the stimuli differed in second formant frequency during the 25 ms transition of consonant to vowel. It is to be noted that a 25 ms transition is abnormally short; a more typical transition duration in conversational speech occurs over a 45-50 ms period (Gay, 1970). The second stimulus pair was constructed by lengthening the duration of the second formant spectral change so that it occurred over 50 ms, a typical duration for a /ba/ or /ga/. The third pair was made by further lengthening the transition to 150 ms, yielding the percept of bisyllabic vowel sequences /ba/ and /ga/. It is important to note that transitions as long as 150 ms rarely, if ever, occur in mature syllables but have been reported in the precanonical babbling vocalizations of infants (Olmer, in press; Olmer et al., 1983). Thus, three stimulus pairs were constructed which differed only with respect to the duration of the second formant transition from consonant to following vowel. Of the three stimulus pairs, only the 50 ms /ba/ vs. /ga/ pair was acoustically motivated by normal articulatory gestures.

Four adults who participated in discrimination and identification testing of the stimulus pairs had no difficulty correctly identifying each of the six stimuli. Discrimination scores were 100% for all adult perceivers.

Stimulus pairs were constructed: one for training and three for testing. Stimuli were recorded with each member of the pair on a different channel of a two-track audiotape. The stimulus on a single channel was separated by an interstimulus interval (ISI) of 500 ms. Computer control permitted switching from one member of the pair to the other without clipping stimuli.

Apparatus

The experimental site consisted of a double-walled, sound-attenuated booth equipped with an HP4100 speaker, four visual reinforcers (housed in a dark plastic box), and a response box allowing communication with the computer. The adjoining control room housed high fidelity playback and amplification equipment and a DEC 11/23 laboratory computer for controlling stimulus conditions and reinforcement and for recording and analyzing data.

Procedure

The Visually Reinforced Inling Speech Discrimination (VRISD) paradigm (as revised by Eilers & Gavin, 1981), in which infants are conditioned to turn their heads to a change in a repeating background auditory stimulus, was used to assess speech discrimination in both the DS and normally-developing infants. During this procedure, the infant was seated on an adult's lap in the booth. An experimenter attempted to keep the infant in a midline orientation by manipulating a set of quiet toys. In all cases the adult holding the infant and the experimenter listened to masking music over headphones throughout the
The session began with the continuous presentation at 65 dB SPL of the training stimulus on one tape channel (e.g., /bit/). This background stimulus (Sb) remained on throughout the training, interrupted by the presentations of the change stimulus (Sd) on the other tape channel (e.g., /bit/).

The experimenter was responsible for monitoring the infant's behavior, initiating trials, and indicating head-turn responses. Each test session began with a shaping phase during which the background stimulus (Sb) was changed to the contrasting stimulus (Sd) for approximately six seconds at an intensity 12 dB greater than the background level. If the infant was observed to turn toward the speaker during Sb, the experimenter activated the reinforcer by depressing a button on the response box. If the infant did not turn on the first trial, the reinforcer was activated toward the end of the 6-sec period of the second trial. After a few trials most infants turned at the presentation of the louder Sb. After two consecutive trials in which the infants responded correctly within the Sb time window, the intensity of Sb was reduced in 4-dB steps until the infant responded correctly on two consecutive trials at each intensity level. At this point matched intensity (65 dB) change trials were introduced.

During the matched intensity phase, the change from Sb to Sd occurred on one-half of the trials (change trials). On the remaining trials (control trials), no change from Sb to Sd occurred. During control trials, head turns were recorded but no reinforcement was given. These no-change trials served to control for the infant's spontaneous, random head-turning. On every fifth trial a probe was presented in which a change from Sb to Sd was accompanied by a 4-dB increment. These probes were included to maintain the infant's interest throughout the session, but head turns during these intensity-cued trials were not included in the analysis. The order of presentation of the change and control trials was pseudorandom with three possible combinations of change and control trials per block of ten trials: six change and four control, five change and five control, or four change and six control. The experimenter initiated a trial when the infant was engaged at midline. The pseudorandom order of trial presentation coupled with the absence of auditory information during testing prevented the experimenter from knowing the nature of the test trial. Hence all observations of head-turning behavior were made free of experimenter bias.

Infants were trained to a criterion of nine out of ten correct successive equal-intensity test trials with the stimulus pair /bit/ vs. /bit/. This generally required between one-and-a-half and two-and-a-half years of age (mean age 20 months) to be able to meet the training criteria for participation in this study of speech discrimination. DS infants met the criterion in an average of 1.9 sessions which was comparable to the average of 1.5 sessions for normally-developing infants (mean age 7.2 months). Performance for each infant on the test pairs was blocked into three groups of 10 test trials each and expressed as a discriminative index (DI) score for each trial block. The DI score is the number of hits minus the number of false positives divided by the number of change trials (see Morse, Eilers, & Gavin, 1982, for a discussion of DI scores in VRISD testing). These scores were subjected to (1) Z-tests of means to determine whether performance on a contrast differed from chance and (2) ANOVA to assess differences in performance between groups and contrasts.

Table 1 presents the mean DI scores, averaged across test blocks, for each of the stimulus contrasts within the different subject groups. Since each contrast was tested with 9 subjects each receiving 30 trials, the DI scores for the individual contrasts were tested against a population mean of 0, a population standard deviation of .198, and a standard error of .066 (Morse et al., 1982). The
results of the Z-tests of means are also included in Table 1. These tests indicated that DI scores were significantly greater than chance for both DS and normally-developing infants on all three pairs of test stimuli. Inspection of Figure 1 indicates that for both groups of infants performance on the normal duration (50 ms) contrast was better than performance on the abnormally fast (25 ms) and the abnormally slow (150 ms) contrasts, though the trend is more pronounced for the normally-developing infants.

**Down Syndrome**

<table>
<thead>
<tr>
<th>Pair</th>
<th>Mean DI</th>
<th>t-test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ba/-/gau/</td>
<td>.22</td>
<td>3.30*</td>
</tr>
<tr>
<td>/ba/-/gau/</td>
<td>.31</td>
<td>4.76*</td>
</tr>
<tr>
<td>/au/-/iou/</td>
<td>.28</td>
<td>4.19*</td>
</tr>
</tbody>
</table>

**Normal**

<table>
<thead>
<tr>
<th>Pair</th>
<th>Mean DI</th>
<th>t-test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ba/-/gau/</td>
<td>.22</td>
<td>3.38*</td>
</tr>
<tr>
<td>/ba/-/gau/</td>
<td>.45</td>
<td>6.75*</td>
</tr>
<tr>
<td>/au/-/iou/</td>
<td>.41</td>
<td>6.25*</td>
</tr>
</tbody>
</table>

* p < .01

**Figure 1:** Mean DI scores and t-test statistics for DS and normally-developing infants.

A split-plot ANOVA with one between-subjects factor (normal vs. DS) and two within-subjects factors (stimulus contrast and trial block) was performed on the DI scores. This analysis indicated a significant effect for contrast, F(2,35) = 4.34, p < .02. Post hoc analyses (Duncan's multiple-range test) indicated that the 50 ms and 150 ms pairs were discriminated significantly better than the fast duration (25 ms) pair, F(1,16) = 4.34, p < .09. The main effect for Subject Group approached significance (F(1,16) = 4.34, p < .09). The main effect for Trial Block was not significant, nor were any interactions. Post hoc analyses revealed that performance of the DS and normal groups differed only on the /ba/-/gau/ pair (p < .03). Two additional ANOVAs, one considering only the DS group and the other only the Normal group were performed with two within-subjects variables, Contrasts and Trial Blocks. These analyses indicated that there were no significant differences between contrasts for DS subjects. The normal subject analysis, however, yielded a significant main effect for Contrast (F = 4.88, p < .01). Post hoc analyses indicated that both /ba/-/gau/ and /au/-/iou/ differed significantly from the /ba/-/gau/ pair.

**Discussion**

The perceiver's ability to process speech information efficiently is limited by temporal constraints. Since a large proportion of the speech code is based upon subtle spectral differences which occur in rapid time frames, a failure to perceive rapid spectral information has far-reaching consequences. Recent research has suggested that some children with language delay have specific temporal processing deficits that could have serious effects on their perception of speech. The current study was designed to explore further the hypothesis that language delayed DS infants also have deficits in temporal processing that would affect their ability to perceive rapidly presented spectral information in speech-like stimuli. Accordingly, both DS and normal infants were presented with three syllable pairs in a discrimination study. One of these syllable pairs involved a formant transition duration (50 ms) common in speech. The other two pairs contained transition durations not commonly found in natural speech, a short transition duration (25 ms) in one pair and a long transition duration (150 ms) in the other.

As expected, both DS and normally-developing children demonstrated the lowest DI scores for the abnormally fast transition duration pair. Both infant groups had relative difficulty processing spectral information presented more rapidly than is typical of stop consonant-vowel syllables. Although there was no significant interaction between infant group and transition duration, the two groups showed a different pattern of performance as a function of the test contrasts. As can be seen in Figure 1, a significant improvement in normal infant performance occurred between the 25- and 50-ms transition pairs, whereas DS infant performance was relatively flat. This difference in the discrimination functions of normal and DS Infants is further suggestive of a temporal processing deficit in DS. DS children appear to differ from normals, primarily in processing transitions in time frames commonly found for stop consonants. Work with a broader range of stimuli will be needed to confirm this hypothesis.

**References**


Schellen, I., & Leamer, Michigan: University Microfilms.


