INTRODUCTION

In contrast to the large body of data on infants' discrimination of segmental speech contrasts (reviewed by Aslin and Pisoni, 1980; Eilers, 1980; and Trehub, 1979), little is known about suprasegmental influences such as stress on infant speech processing. For adult speakers of many languages, linguistic stress reflects variation in fundamental frequency, duration, and intensity (Fry, 1955; Lehiste, 1976), although the relative importance of these factors is complex and has not been fully disentangled (Lieberman, 1960; Nakatani and Aston, 1983). To facilitate future work on the role of stress in infant speech perception, psychoacoustical information on discrimination thresholds and sensory scaling is required.

Research has shown that infants discriminate the position of stressed syllables cued by a multiple change in fundamental frequency, duration, and intensity (Spring and Dale, 1977). Preliminary work on the discrimination of a single stress correlate, duration, indicated that infants were sensitive to relatively small increments (Spring and Dale, 1977), although no attempt was made to scale infants' duration increment thresholds. More recent work has detailed discrimination of a final-syllable vowel-duration increment in one-, two-, and three-syllable contexts (Eilers et al., 1984a). This work represented the first step in scaling the acoustic parameters of a set of stimuli to be employed in future studies on the influence of stress in infant speech processing. Towards such a goal, it is also necessary to scale intensity and fundamental frequency. The present investigation was concerned with infants' sensitivity to an intensity change within the same stimuli employed in an earlier vowel duration study (Eilers et al., 1984a).

Previous research (as reviewed by Schneider et al., 1979) has shown that infants can discriminate intensity differences very early in life, but few reports have attempted to estimate infant intensity-difference thresholds. Infants' thresholds for increments in the amplitude of broadband noise fall within the range from 2 to 5 dB as estimated by a two-alternative, forced-choice procedure (Schneider et al., 1983). Similar results with a turn, no-turn procedure with limited stimulus duration indicate individual infant increment thresholds range from 2 to 12 dB for a 1-kHz tone or a vowel /a/ (Spring and Aslin, 1983).

These earlier threshold studies estimate the auditory capacities of infants in certain listening situations. It is not clear that these same psychophysical limitations apply to the intensity differences associated with stress changes in multisyllabic sequences. Such variations result in a difference in peak intensity within a word rather than a total intensity change across a word. Accordingly, the role of intensity in linguistic stress should be considered by assessing discrimination of intensity changes of individual syllables in multisyllabic contexts.

The present experiment sought to determine infant sensitivity for peak intensity changes on individual syllables within either a two- or three-syllable wordlike unit. This peak intensity increment was positioned in the central portion of the final syllable /mad/, since intensity changes that cue stress typically are centered within vowel nuclei (Lehiste, 1976). Infants were tested for discrimination of synthetic stimuli with 2-, 4-, and 6-dB increments in the peak intensity of the final syllable of either a /samad/ or a /masamad/ stimulus "word."

I. METHOD

A. Subjects

Thirty-three full term infants, 5 to 11 months of age, were solicited from English-speaking homes for the study. Infants were excluded from the final subject sample because they failed the training criteria after three visits (n = 8), had conductive hearing loss as assessed by tympanography and Visual Reinforcement Audiometry (n = 2), or could not return for the necessary number of visits (n = 5). The infants in the final subject sample (n = 18) met inclusion criteria of normal appearance of developmental milestones and normal hearing. The mean subject age at the start of testing for the final sample of 18 infants was 7.7 months. Eleven subjects were male.
B. Apparatus and stimuli

The experimental site consisted of a double-walled, sound-attenuated booth. An adjoining control room housed high-fidelity playback and amplification equipment, and a DEC 11/23 laboratory computer to control stimuli and reinforcement conditions and to record responses. The experimental booth was equipped with an HPM 100 speaker, four visual reinforcers housed in a dark Plexiglas box, and a response box for communication with the computer.

The training stimulus pair consisted of five tokens each of the syllables /sa/ and /da/, produced by a male phonetician and matched pairwise for overall duration, overall amplitude, peak amplitude, mean fundamental frequency, and peak fundamental frequency.

The experimental stimuli consisted of two- and three-syllable stimuli, /samad/ and /masamad/, synthesized using the Klatt (1980) software synthesis routines. Initially, a three-syllable /masamad/ was synthesized to serve as the basis for all other stimulus generation. This stimulus began with a 55-ms /m/ murmur, a 45-ms transition, and a 100-ms steady-state vowel /a/. The steady-state /a/ had an F of 745 Hz (BW = 130), an F2 of 900 Hz (BW = 70), and an F3 of 2400 Hz (BW = 160) in all three syllables. After a 50-ms transition, the 60-ms /s/ friction began, followed by a 45-ms transition into a 100-ms vowel /a/. The second /m/ murmur (F1 = 275, F2 = 900, F3 = 2200), which followed, was 100 ms in duration with a 45-ms transition into the steady-state portion of the vowel. The base duration stimulus had 200 ms of steady-state vowel followed by a 55-ms transition providing a 300-ms vowel. The final consonant, /d/, had an F1 of 150, F2 of 1700, F3 of 2600), had a closure duration of 115 ms followed by an approximate 20-ms burst. The burst sequence also included a 25-ms transition to schwa formant values. During the closure, the amplitude of voicing (AV) dipped to 15 dB below the maximum and was raised by 10 dB during the burst before falling off to 0. Amplitude of friction (AF) and amplitude of aspiration (AH) were also raised during the burst to values of — 10 and — 12 dB re: the maximum AV value, respectively, before falling off. The amplitude of the third through sixth formants during the burst were — 13, 0, 2, and 0 dB re: the maximum AV value. The fundamental frequency contour on the final syllable involved a rising—falling pattern, starting from 100 Hz during the /m/, reaching 150 Hz midway through the vowel, and dropping to 98 Hz at the end of the burst. This /masamad/ had a total duration of 1045 ms and was identical to the base stimulus employed by Eilers et al. (1984a).

The third syllable, /mad/, was synthesized with a relatively flat intensity contour by fixing the amplitude within a 385-ms critical period that started 15 ms after the start of the /m/ and ended at the completion of the transition preceding the final consonant /d/. This base intensity stimulus, denoted /masamad/, was contrasted with stimuli containing either 2-, 4-, or 6-dB increments during this critical period. Each 2-dB increment was added across a 5-ms ramp. Thus the 6-dB stimulus, denoted /masamad + 6/, was synthesized with 2-dB increments every 5 ms after the start of the critical period until the intensity was 6 dB higher than the base level. The intensity then remained at this level until 15 ms prior to the end of the critical period, at which time the intensity decreased 2 dB every 5 ms until it again reached the base level. The 2- and 4-dB increment contrasts, /masamad + 2/ and /masamad + 4/ were synthesized in a similar manner. To provide comparable /samad/ stimuli, the first 248 ms of each of the /masamad/ stimuli was trimmed. Six stimulus contrasts were produced: three /samad/ and three /masamad/ pairs. Within each pair the base stimulus was contrasted with one of the stimuli containing an intensity increment. These testing pairs were recorded on individual audiocassettes, with a constant interstimulus interval (ISI) of 650 ms. The training tape, /sa/ vs /da/, was recorded with a constant ISI of 500 ms. Computer control permitted switching between channels without interrupting stimuli.

For calibration purposes, each test recording was preceded by 60 s of a 200-ms, steady-state vowel /a/, identical to the vowel in the base /mad/. This item was synthesized at the output level used for the test stimuli and was recorded at equal amplitude on the two tracks of each tape. Prior to each testing session, this vowel token was used to equate the amplitude of the two tape channels.

For comparison purposes, the six intensity-increment tapes were presented to five naive adults in both discrimination and identification tests. Testing with the VRISD procedure indicated that all adults discriminated the 2-, 4-, and 6-dB increments in intensity without error. In a subsequent test, these same stimuli and the two control (nonincrement) stimuli were presented individually for identification of the presence and location of stressed syllables. All stimuli with increments were identified as having a stressed final syllable, whereas stimuli without increments were judged to have equally stressed syllables. These tests indicate that the increments employed in the infant testing were sufficient for adults to perceive syllables as stressed.

C. Procedure

The present study was based on the Visually Reinforced Infant Speech Discrimination (VRISD) paradigm (Eilers et al., 1981) in which infants are conditioned to turn their heads to a change in a repeating background auditory stimulus. The infant was seated on a parent's lap in the booth so that the speaker was approximately 30° to the right of the infant's midline. An experimenter, seated to the infant's left, attempted to keep the infant at midline orientation by manipulating a set of quiet toys. During the entire session, both the experimenter and the parent listened to masking music through headphones. The experimenter also wore earplugs as an added precaution to mask the test stimuli.

At the beginning of a session, the speaker carried a continuously repeating background stimulus at 60 dB (C scale), measured from a location directly in front of the infant's position. A trial was initiated when the infant was facing the experimenter, at which time the experimenter pressed a button to begin an approximately 6-s observation period. Initially, this observation period always involved a change from the background stimulus to the repeated presentation of a second, contrasting stimulus on the other channel of the tape recorder. If the infant turned towards the speaker during this observation period, the experimenter pressed a button on the
response box and one of the toys in the box above the speaker was illuminated and activated for 4 s. The completion of the observation period ended a trial and the repeating background stimulus was again presented.

Each session began with a shaping phase in which the contrasting stimulus was presented at an intensity of 12 dB greater than the background stimulus. If the infant did not turn on the first trial, then one of the toys was activated towards the end of the next 6-s period. Most infants turned to the presentation of the contrasting stimulus with only a few presentations. When the infant responded correctly on two consecutive trials, the contrasting stimulus was reduced by 4 dB. The shaping phase then continued, with successive 4-dB decreases in the contrast stimulus after two consecutive correct trials, until successful testing with a 4-dB difference between stimuli was completed. The test phase with matched intensity levels (60 dB) then began.

During the test phase, the observation period did not always involve a change in the auditory stimulus. On control trials, the auditory background remained on and headturns were recorded, but without the toy reinforcement. These no-change trials served to control for the spontaneous, random headturning. The order of presentation of the change and control trials was pseudorandom with three possible combinations of change and control trials for every block of ten trials: six change and four control, five change and five control, or four change and six control. The test phase consisted of 30 trials, approximately half of which were control trials. In addition, after every five test trials, a change trial was presented 4 dB above the testing intensity level. These presentations were probe trials to maintain the infant's interest throughout the session. Headturns during these intensity-cued probe trials were not included in any of the analyses.

Infants returned to the laboratory for testing an average of one session per week. The initial sessions involved training to a criterion of nine out of ten consecutively correct equal-intensity test trials with the stimulus pair /da/ vs /sa/. Note that either a headturn for a change trial or no response for a control trial is a correct trial. Upon completion of this criterion, which generally required one to three sessions, the next three sessions involved testing with specified experimental contrasts. For this testing, the member of a stimulus pair with an increment in the intensity of /samad/ was always used as the contrasting stimulus. For training, the use of /da/ or /sa/ as the background stimulus was counterbalanced across subjects.

One group of nine infants was tested with the three bisyllabic /samad/ intensity contrasts, while the second group received the three trisyllabic /masamad/ contrasts. Subjects were tested with a different contrast each week. The contrasts were tested in a counterbalanced order across subjects.

II. RESULTS

For each subject, the number of headturns on change trials (hits) and control trials (false alarms) was blocked into three groups of ten test trials each. Performance for each trial block was expressed by a discriminative index (DI) score representing the number of hits minus the number of false alarms divided by the number of change trials (see Morse et al., 1982 and Eilers et al., 1984b, for a comparison of the relative advantages of DI and percent correct scores in infant discrimination testing). These scores were subjected to both z tests of means to determine whether performance on a contrast differed from chance, and ANOVA to assess differences in performance for the different contrasts.

Table I presents the mean DI scores averaged across test blocks for the three intensity increments and the two different syllable-number groups. For comparison purposes, the percent correct (PC) scores are also included in Table I, although results are discussed only for the DI scores. Inspection of Table I indicates that discrimination scores increased as the intensity increment increased. For all intensity increments, mean DI scores are higher for the two-syllable group. Since each contrast was tested with nine subjects receiving 30 trials, DI scores for the individual contrasts were tested against a population mean of 0 with a standard error of 0.066 (Morse et al., 1982). The z-statistic tests are included in Table I. These tests indicate that DI scores were significantly above chance performance for all intensity increments.

A split-plot ANOVA with one between-subjects factor (syllable number) and two within-subjects factors (intensity increment and trial block) was performed on these DI scores. This analysis indicated a significant effect for both increment, F(2, 32) = 11.5, p < 0.001 and syllable number, F(1,16) = 10.1, p < 0.01. The interaction between these two factors was not significant, F(2,32) = 1.45, p > 0.2. Scheffe's post hoc analyses revealed that discrimination of the 2-dB increment was significantly different than performance for

**TABLE I.** Mean, z statistic, and t statistic for both DI and PC scores as a function of the final-syllable intensity increment (2, 4, and 6 dB) for the 2- and 3-syllable stimuli.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/samad/ vs /samad+/</td>
<td>/masamad/ vs /masamad+/</td>
</tr>
<tr>
<td>/samad/</td>
<td>/masamad/</td>
</tr>
<tr>
<td>/samad+/</td>
<td>/masamad+/</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>DI scores</th>
<th>PC scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t</td>
<td>z</td>
<td>t</td>
</tr>
<tr>
<td>Group 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/samad/ vs /samad+/</td>
<td>0.209</td>
<td>(3.58*)</td>
<td>3.17*</td>
</tr>
<tr>
<td>/samad/</td>
<td>0.459</td>
<td>(6.15*)</td>
<td>6.94*</td>
</tr>
<tr>
<td>/samad+/</td>
<td>0.606</td>
<td>(9.74*)</td>
<td>9.18*</td>
</tr>
<tr>
<td>Group 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/masamad/ vs /masamad+/</td>
<td>0.155</td>
<td>(2.80*)</td>
<td>2.35*</td>
</tr>
<tr>
<td>/masamad/</td>
<td>0.260</td>
<td>(3.71*)</td>
<td>3.94*</td>
</tr>
<tr>
<td>/masamad+/</td>
<td>0.346</td>
<td>(6.24*)</td>
<td>5.24*</td>
</tr>
</tbody>
</table>

* p < 0.05.

**p < 0.01.
either the 4-dB ($p < 0.05$) or the 6-dB ($p < 0.01$) increment contrasts. No significant difference was found between the 4-dB vs 6-dB increment conditions. No other main effects or interactions were significant.

III. DISCUSSION

The results for a final-syllable, peak intensity increment in a multisyllabic context indicate an orderly relation between the increment (2, 4, or 6 dB) and infants' discrimination scores. As the intensity increment was increased, the obtained DI score increased. Although DI scores for the 2-dB increment were relatively low, performance was still above chance expectation. In pilot testing with the 2-dB increment stimuli, adult subjects performed at 100% correct levels, suggesting that adult threshold values would be less than the infant values. Adult-infant differences for intensity and frequency differences have been reported in earlier studies by Schneider et al. (1983) and Sinnott and Aslin (1983). Nevertheless, the results indicate infants to be quite sensitive to intensity increments.

A value within the range of 2 to 6 dB as an estimate of an infant increment threshold is comparable to the threshold estimates reported by Schneider et al. (1983) for a broadband noise increment and by Sinnott and Aslin (1983) for an isolated vowel /a/. It is important to note that the increment in the present experiment was in peak intensity, resulting in a change in the relative shape of the power envelope. Thus it is difficult to compare this study with other studies employing absolute differences in total intensity level. These results do suggest, however, that infants are able to discriminate peak intensity variations in multisyllabic stimuli comparable to those used to cue certain stress contrasts in natural languages. English stress contrasts for the initial syllable in word pairs such as OBJECT/obJECT yield peak intensity differences of about 5 to 6 dB (Fry, 1955). In other cases, English stress contrasts do not seem to correlate simply with intensity (Nakatani and Aston, 1983). The results of the present study, then, suggest that infants' discrimination capabilities for vowel intensities could be useful for discrimination of some linguistic stress contrasts.

Infant subjects were more sensitive to an intensity difference in the two-syllable than in the three-syllable context. A previous study of infants' discrimination of final-syllable vowel lengthening found no effects of syllable number (Eilers et al., 1984a), although other research has shown that increases in syllable number can decrease infants' sensitivity for certain phonetic contrasts (Trehub, 1973, 1976). Other research with phonetic contrasts has found no effect of syllable number on infant discrimination (Jusczyk and Thompson, 1978). In the present experiment, it would appear that infants are better able to discriminate an intensity variation carried within a greater proportion of the total stimulus. The exact explanation for syllable-number effects in infant intensity discrimination is unclear, although it is possible that the extra /ma/ syllable may increase the information-processing load of the discrimination task. At the present time, no clear pattern of data has emerged to support a general conclusion about syllable-number effects in infant speech processing.

Infant sensitivity for a final-syllable, peak intensity increment can be compared directly with results for a final-syllable, vowel-duration increment (Eilers et al., 1984a). Infant DI scores for a 2- or 4-dB increment in /masamad/ stimuli and for a 2-dB increment in /samad/ stimuli fall within the range of discrimination performance for 100-, 200-, and 300-ms increases in a base vowel duration of 300 ms. Infants were more sensitive, however, to a peak intensity increment of 6 dB than to a vowel-duration increase of 100%. Results from these and other comparisons can be used to "equalize" basic acoustic parameters in future research with both speech and nonspeech stimuli. Such psychophysical work involving intensity, as well as duration and frequency is a necessary prerequisite to understanding the relative influence of suprasegmental factors in infant speech perception.

ACKNOWLEDGMENTS

This work was supported by NIH grant IR01 HD12701 and by the Mailman Foundation. We would like to thank the infants and parents for their participation. Our special thanks to Philip Morse, Cliff Gillman, Don Childers, and the staffs of their respective laboratories for assistance with the software for stimulus preparation. Reprints may be obtained from Dale Bull, Mailman Center for Child Development, University of Miami, P. O. Box 016820, Miami, FL 33101.

Although $z$ tests appear to be the most appropriate statistic since both the population mean and standard deviation are known (see Hays, 1966), $t$ tests may also be performed to assess differences from chance performance. The $t$ test results are presented in parentheses in Table I next to the appropriate scoring method.

A value within the range of 2 to 6 dB as an estimate of an infant increment threshold is comparable to the threshold estimates reported by Schneider et al. (1983) for a broadband noise increment and by Sinnott and Aslin (1983) for an isolated vowel /a/. It is important to note that the increment in the present experiment was in peak intensity, resulting in a change in the relative shape of the power envelope. Thus it is difficult to compare this study with other studies employing absolute differences in total intensity level. These results do suggest, however, that infants are able to discriminate peak intensity variations in multisyllabic stimuli comparable to those used to cue certain stress contrasts in natural languages. English stress contrasts for the initial syllable in word pairs such as OBJECT/obJECT yield peak intensity differences of about 5 to 6 dB (Fry, 1955). In other cases, English stress contrasts do not seem to correlate simply with intensity (Nakatani and Aston, 1983). The results of the present study, then, suggest that infants' discrimination capabilities for vowel intensities could be useful for discrimination of some linguistic stress contrasts.

Infant subjects were more sensitive to an intensity difference in the two-syllable than in the three-syllable context. A previous study of infants' discrimination of final-syllable vowel lengthening found no effects of syllable number (Eilers et al., 1984a), although other research has shown that increases in syllable number can decrease infants' sensitivity for certain phonetic contrasts (Trehub, 1973, 1976). Other research with phonetic contrasts has found no effect of syllable number on infant discrimination (Jusczyk and Thompson, 1978). In the present experiment, it would appear that infants are better able to discriminate an intensity variation carried within a greater proportion of the total stimulus. The exact explanation for syllable-number effects in infant intensity discrimination is unclear, although it is possible that the extra /ma/ syllable may increase the information-processing load of the discrimination task. At the present time, no clear pattern of data has emerged to support a general conclusion about syllable-number effects in infant speech processing.

Although $z$ tests appear to be the most appropriate statistic since both the population mean and standard deviation are known (see Hays, 1966), $t$ tests may also be performed to assess differences from chance performance. The $t$ test results are presented in parentheses in Table I next to the appropriate scoring method.


