Organizational Processes in Infants' Perception of Auditory Patterns

Sandra E. Trehub, Leigh A. Thorpe, and Barbara A. Morrongiello

University of Toronto

TREHUB, SANDRA E.; THORPE, LEIGH A.; and MORRONGIELLO, BARBARA A. Organizational Processes in Infants' Perception of Auditory Patterns. CHILD DEVELOPMENT, 1987, 58, 741-749. Infants 9–11 months of age were tested for their discrimination of changes in the melodic contour (direction of successive pitch changes) of brief melodies in the context of discernible variations in key (different absolute frequencies, same intervals) or interval size (different absolute frequencies and frequency ratios, same contour). Infants detected contour changes in both variable contexts, suggesting that they categorize sequences of sounds on the basis of global, relational properties such as melodic contour. Implications of such processing strategies for infants' perception of running speech are considered.

Under many listening conditions, adults ignore acoustic variations among exemplars of specific syllables, words, or other familiar sounds. There is some indication that infants do likewise. For example, infants 6 months of age perceive the similarity of speech syllables despite variations in voice quality and pitch contour (Hillenbrand, 1983; Kuhl, 1979, 1983) and the similarity of sounds produced by same-sex speakers (Miller, 1983; Miller, Younger, & Morse, 1982). Infants 6 months of age and beyond also respond equivalently to complex tones that vary discernibly in spectral composition, intensity, or duration (Clarkson & Clifton, 1985; Endman, 1986). In these investigations, infants' perception of similarity was established by demonstrating discrimination between sets of sounds in the context of acoustic variation within each set. If infants' ability in this regard were restricted to the categorization of single sounds, as in the aforementioned studies, then its practical significance would be limited.

Adults also ignore certain classes of variation in sequences of sounds. For example, they recognize sequences of tones (i.e., melodies) over transformations that alter the absolute frequency of individual tones but preserve the precise frequency relations between tones (i.e., transpositions or key changes). In such cases, it may be easier to remember information about the frequency relations between tones than about the absolute frequencies (Deutsch, 1969).

There is some suggestion that infants' perception of melodies is similarly holistic, in the sense that global properties are retained after exposure to specific exemplars. For example, following relatively limited exposure to a brief tone sequence or melody, infants fail to respond to an altered melody (i.e., different absolute frequencies) that preserves the intervals (i.e., ratios of adjacent frequencies) or melodic contour (i.e., direction of successive frequency changes without regard to the size of such change) of the original but do respond to an altered melody that changes the contour or frequency range of the original melody (Chang & Trehub, 1977a; Trehub, Bull, & Thorpe, 1984, Experiment 2). One implication of the failure to respond to same-contour and same-interval melodies is that infants encode and retain global information about the original melody at the expense of local information about absolute frequency.

This failure to respond may also reflect infants' perception of the similarity between standard and comparison melodies, which shared the same pattern of intervals or contour. Since the foregoing studies of infant melody perception employed a single stan-
standard and comparison melody in each condition, one could only establish whether infants detected differences between them, not whether they perceived similarities. Infants' perception of such similarities could be demonstrated by their discrimination between sets of standard and comparison melodies. If infants could categorize discernible tone sequences on the basis of common intervals or contour, this would not only confirm infants' effective processing of multi-element signals but would also provide insight into their internal representation of complex auditory information.

Since the terms melodic contour, interval, and key derive from music rather than psychology, some clarification may be in order (see Dowling & Harwood, 1986, for a detailed description of these concepts). Individual notes of a melody can rise in pitch relative to the immediately preceding note (+), fall in pitch (−), or stay the same (0). The overall pattern of such changes in pitch over the course of a melody defines its contour. What is critical here is the direction of pitch change rather than the extent of such change. Thus, for example, a three-note passage that rises in pitch from first to second note and falls from second to third (i.e., contour of + −) is considered to have the same contour as another three-note sequences that do likewise, regardless of the absolute frequency of individual notes. Intervals, on the other hand, define the pitch relations between adjacent notes, which can be expressed nontechnically as the ratio of these adjacent pitches, since the frequency scale (Hertz, or Hz) is logarithmic. Two sequences of notes that have the same contour may or may not have the same intervals. If they do, they are said to be transpositions, and involve either raising or lowering the overall sequence to another musical key.

The key of a melody simply implies adherence to the component notes of one of the major or minor scales, which, in turn, are defined by a specified sequence of whole and half steps. A musical passage is said to be in a particular key, that key generally being named for the note on which the passage begins and ends (or its dominant influence). All major keys are exactly alike, except for the starting frequency, which is to say that they are all transpositions of one another. The same holds for minor keys. Nevertheless, some keys are considered to be more closely related than others (see Dowling, 1982), with "near" keys having many common notes in their scales (e.g., C major and G major share every note but one) and "far" keys having few notes in common (e.g., C major and B major share two notes only).

In the present investigation, infants were required to discriminate between five-tone melodies with contrasting contour in the context of variations that altered absolute frequency but preserved interval size (i.e., transpositions) or variations that altered absolute frequency and interval size but preserved contour. Specifically, infants were presented with a repeating background of variable tone sequences characterized by common intervals or contour and were rewarded for responding only when one of a set of sequences with contrasting contour replaced the background sequence. In order to perform such a task, infants must perceive the global frequency configuration of the melodies, that is, they must extract either the pattern of interval relations or the contour of each sequence. They must then categorize the sequences on the basis of similar interval relations or contour. Finally, they must respond only to sequences that have attributes inconsistent with their previous categorization.

In the case of adults, there is evidence that contour information is easily encoded (Bartlett & Dowling, 1980; Dowling & Fujitani, 1971). In fact, musically untrained listeners are almost as good as trained listeners at recognizing contour (Dowling, 1978). In contrast, interval information is difficult to encode from unfamiliar melodies (Dowling & Fujitani, 1971) and musical training does facilitate its perception (Bartlett & Dowling, 1980; Cuddy & Cohen, 1976). Nevertheless, interval information, once encoded, fades slowly, and there is some suggestion that access to contour information is difficult once interval information has been encoded (Dowling & Bartlett, 1981).

If infants encode the intervals of melodic patterns, then they should not have difficulty responding to contour change in the context of transpositions, since the change of contour (i.e., target change) necessarily involves interval changes, but the irrelevant background variation involves no such interval changes. On the other hand, they should have difficulty responding to contour change in the context of contour-preserving transformations, since both the target and irrelevant variations embody interval changes. If infants only encode the less precise contour information, then both tasks would be essentially the same for them and they should perform equivalently. Finally, if they fail to encode configurational aspects of the melodies, such as intervals or contour, and, instead, encode absolute fre-
quencies, then they should be unable to perform either task since the target and irrelevant variations both incorporate frequency changes.

On the basis of infants' previous discrimination of contour changes in the context of fixed, six-tone melodies (Chang & Trehub, 1977a; Trehub et al., 1984) coupled with their discrimination of speech and nonspeech stimuli in the context of variable exemplars (Clarkson & Clifton, 1985; Endman, 1986; Kuhl, 1979, 1983), we expected successful discrimination of contour changes in one or both variable contexts. On the basis of infants' comparable performance on the detection of transpositions and contour-preserving changes (Trehub et al., 1984) coupled with adults' reported difficulty encoding the intervals of unfamiliar melodies (Dowling, 1978), we expected that infants would encode contour rather than intervals. As a result, we anticipated similar performance in the two variable contexts.

Method

Subjects.—The subjects were healthy, full-term infants from 9 to 11 months of age, as were the subjects of previous research on the discrimination of melodic contour with the operant head-turn procedure (e.g., Trehub et al., 1984). Infants were excluded from the sample if they failed to meet the predetermined training criterion described in the Procedure section (N = 6), or if they failed to complete the 30-trial test session due to crying or fussing (N = 3). The final sample of 54 infants consisted of 29 males and 25 females, with a mean age of 10 months. There were 10 infants in the key-variation condition, 10 in a replication of that condition, 10 in the interval-variation condition, and six in each of four control conditions.

Apparatus.—The experiment was controlled by a microcomputer (Commodore PET, Model 2001), which operated the remaining electronic equipment through a custom-built interface. The stimuli were generated on line by a synthesizer/function generator (Hewlett-Packard 3325A) and presented via one channel of a stereo amplifier (Technics Model SU 7300) over a single loudspeaker (Radio Speakers of Canada, Model C8C). The loudspeaker was positioned above a four-chamber smoked Plexiglas box that contained four different mechanical toys. Stimulus intensity was controlled by the synthesizer and calibrated with a General Radio (Model 1551-C) sound-level meter. Tone onset, offset, and ramping characteristics were controlled through an electronic switch (Med Associates ANL-913). Testing was carried out inside a double-walled, sound-attenuating chamber (Industrial Acoustics Co.).

Stimuli.—There were two experimental conditions in which the standard and contrasting stimuli consisted of sets of melodies with musical key (key-variation condition) or both musical key and interval size (interval-variation condition) as the irrelevant variation. The discrimination target in both conditions was a change in melodic contour. Four control conditions evaluated infants' discrimination of individual exemplars from the background set of each experimental condition. The melodies from each experimental set and from the control conditions are shown in Table 1. Selected examples are illustrated in Figure 1.

The standard sequence of the key-variation condition was a five-tone melody based on the major triad, with frequencies consistent with the well-tempered scale (Bakus, 1969, p. 134). This melody was transposed into each of five keys, which were chosen to be closely related in the musical sense. The contrasting melody of the key-variation condition was composed of the same tones as the standard, reordered to obtain a different melodic contour; this melody was transposed into the same keys as the standard set. The standard and contrasting sequences in the key of C major are illustrated in panel a of Figure 1. The standard stimulus set of the interval-variation condition varied musical key and relative interval size while preserving melodic contour over the set of five exemplars. The contrasting sequences for each key consisted of the identical tones of the standard, reordered as in the key-variation condition to obtain a different melodic contour. A visual analogue of the frequency intervals for several members of the standard and contrasting sets is given in Figure 1. One member of the interval-variation standard set had the same intervals as all the standards of the key-variation condition (panel a). Other exemplars of the interval-variation standard set, along with the corresponding contrasting sequences, are given in panels b, c, and d. Inspection of each row in Figure 1 reveals that both the sequence of intervals and the contour differ for standard and contrasting sequences in each key, but the frequencies comprising each sequence are simply reordered. Inspection of each column in Figure 1 conveys a sense of the variation incorporated in each set of the interval-variation condition.
<table>
<thead>
<tr>
<th>Condition and Key</th>
<th>Background Set</th>
<th>Contrast Set</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key variation:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>233 294 349 294 233</td>
<td>349 294 233 294 233</td>
</tr>
<tr>
<td>F</td>
<td>349 440 523 440 349</td>
<td>523 440 349 440 349</td>
</tr>
<tr>
<td>C</td>
<td>262 330 392 330 262</td>
<td>392 330 262 330 262</td>
</tr>
<tr>
<td>G</td>
<td>392 494 587 494 392</td>
<td>587 494 392 494 392</td>
</tr>
<tr>
<td>D</td>
<td>294 370 440 370 294</td>
<td>440 370 294 370 294</td>
</tr>
<tr>
<td><strong>Interval variation:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>233 349 392 349 233</td>
<td>392 349 233 349 233</td>
</tr>
<tr>
<td>F</td>
<td>349 440 466 440 349</td>
<td>466 440 349 440 349</td>
</tr>
<tr>
<td>C</td>
<td>262 349 392 349 262</td>
<td>392 349 262 349 262</td>
</tr>
<tr>
<td>G</td>
<td>392 440 494 440 392</td>
<td>494 440 392 440 392</td>
</tr>
<tr>
<td>D</td>
<td>294 370 440 370 294</td>
<td>440 370 294 370 294</td>
</tr>
<tr>
<td><strong>Key-variation-control:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest absolute difference</td>
<td>262 330 392 330 262</td>
<td>294 370 440 370 294</td>
</tr>
<tr>
<td>Most related keys</td>
<td>262 330 392 330 262</td>
<td>392 494 587 494 392</td>
</tr>
<tr>
<td><strong>Interval-variation-control:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change to four intervals</td>
<td>262 349 392 349 262</td>
<td>294 370 440 370 294</td>
</tr>
<tr>
<td>Change to two intervals</td>
<td>294 370 440 370 294</td>
<td>349 440 466 440 349</td>
</tr>
</tbody>
</table>
In each experimental condition, the repeating background stimulation consisted of a quasi-random series of tokens of the standard set, with the sequence determined in such a way that each change of key was always made to one of the nearest musical keys, and the same key was never presented twice consecutively. In other words, key changes in the background were a random walk through the set BV, F, C, G, and D. A randomized block of 100 exemplars was generated prior to testing. This order was recycled as necessary to achieve a continuously repeating background for both the training and testing phases. On any particular change trial, the contrasting melody was presented in the key designated for the background melody that it replaced. Thus, the key changes between background and contrasting melodies followed the same constraints as key changes between successive background melodies.

In the control conditions, infants' discrimination of each of two pairs of melodies from the standard set of each experimental condition was assessed (see Table 1). Infants in the key-variation-control condition were tested either with a pair of melodies with the smallest absolute log-frequency difference or with a pair with most related musical key. Exemplars from the key of C and D were chosen to examine log-frequency difference, while those from the keys of C and G were chosen to test the discrimination of transpositions to highly related keys. Infants in the interval-variation-control condition were tested with a pair differing in four intervals or a pair differing in two intervals only. Exemplars from the keys of C and D were used to examine discrimination of same-contour melodies differing in four intervals, while those in the keys of D and F were chosen to examine discrimination with two intervals changed.

The component tones of each sequence were sinusoidal waveforms, 200 msec in duration, with rise and fall times of 30 msec and intertone intervals of 200 msec; thus each melody was 1.8 sec in duration. A 1,400-msec silent interval separated adjacent melodies. Stimulus intensity, measured at the approximate location of the infant's head, averaged 68 dB-C. Ambient noise level, measured at this location, was approximately 48 dB-C (26 dB-A).

Procedure.—During the session, the infant was seated on the parent's lap in one corner of the booth, facing the experimenter. To the infant's left, at an angle of 45 degrees, were the loudspeaker and toy display box. The experimenter and parent both wore headphones carrying music to mask the nature of the trials presented to the infant. The experimenter initiated trials and recorded responses with a small control box interfaced to the computer.

The infant was presented with a continuously repeating stimulus background consisting of the random series of tokens described above (for the variation conditions) or one of the standard melodies (for the control conditions). When the infant was quiet and facing directly ahead, the experimenter initiated a training trial, at which time the contrasting melody was substituted for the standard melody. This contrasting melody was presented once, at an intensity 5 dB higher than the previously heard standard melody, after which the standard stimuli were presented repeatedly, as before. If the infant made a head turn of 45 degrees or greater toward the loudspeaker, the experimenter recorded the turn. If the head turn occurred during the presentation of the contrasting sequence or the following 1,200 msec (i.e., response interval of 3.0 sec), one of the mechanical toys was automatically illuminated and activated for 4 sec. Head turns at other times were not reinforced. If the infant responded correctly on two consecutive trials, the intensity of the
Infants discriminated between all melody pairs tested in the control conditions. For the key-variation-control condition, infants turned on 56% of change trials compared to 14% of no-change trials for the melody pair with smallest absolute frequency difference. A mean $d'$ of 1.26 was obtained for this pair, $t(5) = 4.98, p < .005$. For the melody pair with highly related musical keys, infants turned on 67% of change trials, compared to 32% of no-change trials, with a mean $d'$ of 0.95, $t(5) = 5.06, p < .005$. The difference in mean $d'$ between these two key-variation-control pairs was not significant. The results for the interval-variation-control condition were similar. When the melody pair had four contrasting intervals, infants turned on 54% of change trials compared to 20% of no-change trials, with a mean $d'$ of 1.08, $t(5) = 4.99, p < .005$. When the melody pair had only two contrasting intervals, infants turned on 64% of change trials and 30% of no-change trials, with a mean $d'$ of 1.00, $t(5) = 4.30, p < .005$. Again, performance on the two melody pairs did not differ.

Two groups of infants were tested in the key-variation condition. In the initial sample, infants turned on 43% of change and 19% of no-change trials. The mean $d'$ was 0.87, which was significantly greater than zero, $t(9) = 4.09, p < .005$. These results were replicated with a subsequent sample of infants, who turned on 49% of change and 23% of no-change trials. For this group, the mean $d'$ was 0.77, which was also significantly greater than zero, $t(9) = 4.17, p < .005$. Performance differences between the original and replication samples were not statistically significant. The two groups were pooled, yielding response rates of 46% and 21% on change and no-change trials, respectively, and an overall mean $d'$ of 0.82, $t(19) = 5.96, p < .0005$. Thus, infants were successful at discriminating the change in melodic contour in the context of irrelevant variations in musical key. For the interval-variation condition, infants turned on 49% of change compared to 22% of no-change trials; mean $d'$ was 0.87, which was significantly greater than zero, $t(9) = 4.32, p < .005$. A two-tailed $t$ test comparing the interval-variation condition and the pooled key-variation condition revealed no differences in performance.

Discussion

Infants detected changes in melodic contour in the context of discriminable variations in absolute frequency or interval size. Their detection of contour changes replicates previous evidence obtained with tonal (Trehub et al., 1984) and atonal (Chang & Trehub, 1977a)
melodies, and their ability to ignore discernible variations in melodies extends previous research on the perception of single sounds in variable contexts. Infants go beyond categorizing single sounds based on phonetic identity (Kuhl, 1979, 1983), sex of speaker (Miller, 1983; Miller et al., 1982), or fundamental frequency, harmonic structure, intensity, and duration (Clarkson & Clifton, 1985; Endman, 1986) to categorize sequences of sounds based on global, relational properties such as melodic contour (i.e., pattern of ups and downs in frequency). It is important to emphasize that this is the first investigation in which infants were required to discriminate between sets of multi-element sequences on the basis of global patterning as opposed to absolute cues.

The absence of performance differences under conditions of key and key-plus-interval variation implies that infants encoded contour, rather than interval, information. Although this result represents a failure to reject the null hypothesis, our interpretation is supported both by a critical analysis of the present discrimination tasks and by findings in previous studies of infants' and adults' perception of melodies. For an infant who encodes contour, both the key- and interval-variation tasks would be identical, since both types of variation conserve contour. Moreover, the task of responding to changes in contour in a variable context should be considerably more difficult than previous contour discrimination tasks with a fixed background (Trehub et al., 1984; Trehub, Thorpe, & Morrongiello, 1985). An infant who encodes intervals should have no difficulty with the key-variation task, which conserves intervals in the background set but not the contrasting set. In contrast, the interval-variation task, with its variable intervals in the background and contrasting sets, would be impossible for such infants. Since both tasks required detection of the same change of contour, it is unlikely that infants in one condition encoded intervals and those in the other encoded contour.

Our interpretation is consistent with adults' acknowledged difficulty of encoding interval information from unfamiliar melodies (Dowling & Fujitani, 1971), their relative ease of encoding contour information (Bartlett & Dowling, 1980; Dowling, 1978), and infants' reported difficulty with interval discrimination (Trehub et al., 1984, Experiment 2). Recent studies in our laboratory further underscore the salience of contour as opposed to interval size by indicating that infants' detection of contour changes in variable two-tone sequences is no better for tones separated by six semitones than for those separated by one semitone (Thorpe, 1986). In the light of the aforementioned considerations, the most parsimonious interpretation of the present findings is that infants categorized the melodies in both conditions on the basis of melodic contour.

It is possible that infants accomplished these tasks by attending to pitch relations between the first two tones of each sequence rather than the entire sequence. This is unlikely, however, in view of previous evidence that infants can detect changes in any position of five- or six-tone melodies coupled with the absence of enhanced performance for tones early in such sequences (Trehub et al., 1985; Trehub, Cohen, Thorpe, & Morrongiello, 1986). Moreover, there are clear indications that adults have more difficulty encoding information from two-tone sequences than from longer sequences (e.g., Cuddy & Cohen, 1976).

What organizational processes are implicated in the present demonstration of melody categorization on the basis of contour? To hear the background as repeating six-tone melodies rather than successive single tones, infants must first group the tones on the basis of proximity. There is evidence that infants can accomplish such a temporal perception task as early as 2–5 months of age (Chang & Trehub, 1977b; Demany, McKenzie, & Vurpillot, 1977). They must then extract the configuration of directional frequency changes in each melody (i.e., the contour) from the array of changing absolute frequencies. Such abilities have been documented with infants 6–11 months of age (Trehub et al., 1984, 1985). Finally, infants must group the resulting configurations (i.e., contours) on the basis of similarity. Although infants in this age range can categorize or group single sounds (Clarkson & Clifton, 1985; Endman, 1986; Kuhl, 1979, 1983), only the present findings attest to their grouping of configurations on the basis of similarity. The youngest age at which they can accomplish such configurational grouping remains to be determined, but there is some indication that infants younger than 6 months of age experience difficulty categorizing some aspects of single sounds (e.g., Miller, 1983).

These organizational processes have implications for infants' perception of other auditory sequences such as speech. There has been considerable debate concerning the unit of analysis of speech stimuli, with some favoring phonetic features (Eimas, 1982) and
others, larger and less specialized units like syllables (Bertoncini & Mehler, 1981). For the prelinguistic listener, however, there is no compelling reason to expect segmentation rules comparable to the child or adult. Rather, one might expect that suprasegmental parameters (e.g., intonation) would have processing priority in early life (Crystal, 1973; Lewis, 1951; Lieberman, 1967). There is evidence, for example, that pitch contours play a critical role in infants’ response to mother’s voice (Fernald, 1985; Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978).

The importance of pitch contours is also implied in infants’ ability to reproduce speechlike intonation contours and brief melodies prior to their production of meaningful speech (Papousek & Papousek, 1981). Some investigators have even proposed that “the early appreciation of intonation might provide infants with a primitive way of segmenting an otherwise continuous stream of sound” (Kaplan & Kaplan, 1971, p. 371). Thus, the earliest step in infants’ analysis of running speech may be the extraction of its fundamental frequency configuration, which corresponds roughly to the contour of a musical pattern. This would allow infants to recognize their mother’s voice or other infant-directed speech by its characteristic “tune.” Subsequent developmental changes would include the addition of a further, phonetic level of analysis and, ultimately, the subordination of intonational patterning to phonetic patterning.

Papousek and Papousek (1981) have argued that musical elements in the caretaker’s vocalization attract the infant’s attention, modulate the infant’s state, and help the infant to identify the speaker (e.g., mother) and addressee (himself/herself). Conversely, they note that frequency modulation in the infant’s own vocalization directs the caretaker’s attention by providing information about the infant’s behavioral state and physiological needs. In this way, “musical elements appear to be a rich learning material, no matter whether delivered to the infant in the caretaker’s babytalk or generated by the infant itself” (Papousek & Papousek, 1981, p. 216).

References


Trehub, Thorpe, and Morrongiello 749


