Journal of Experimental Psychology:
Human Perception and Performance
1992, Vol. 18, No. 2, 394-402

0096-1523/92/$3.00

A Comparison of Infants' and Adults' Sensitivity to Western Musical Structure

Laurel J. Trainor and Sandra E. Trehub
Centre for Research in Human Development
University of Toronto, Erindale Campus, Mississauga, Ontario, Canada

Adults (n = 28) and 8-month-old infants (n = 48) listened to repeated transpositions of a 10-note melody exemplifying the rules of Western tonal music. They were tested for their detection of two types of changes to that melody: (a) a 4-semitone change in 1 note that remained within the key and implied dominant harmony (diatonic change) or (b) a 1-semitone change in the same note that went outside the key (nondiatonic change). Adults easily detected the nondiatonic change but had difficulty with the diatonic change. Infants detected both changes equally well, performing better than adults in some circumstances. These findings imply that there are qualitative differences in infants' and adults' processing of musical information.

Recently, it has become clear that musical aspects of speech and tonal patterns (e.g., pitch contours, rhythmic patterning) are particularly salient for infant listeners (see Trehub, 1987, 1989, 1990; Trehub & Trainor, in press). For example, infants 5 to 11 months of age readily discriminate tone sequences, or tunes, differing in pitch contour (up/down/same pattern of pitch changes; Chang & Trehub, 1977; Trehub, Bull, & Thorpe, 1984; Trehub, Thorpe, & Morrongiello, 1985), and they can categorize such sequences on the basis of contour (Ferland & Mendelson, 1989; Thorpe, 1986; Trehub, Thorpe, & Morrongiello, 1987). There is evidence, moreover, of basic grouping processes in infancy, including grouping by similarity (Demany, 1982; Thorpe & Trehub, 1989; Thorpe, Trehub, Morrongiello, & Bull, 1988) and by proximity (Thorpe & Trehub, 1989). Infants are also sensitive to temporal cues for phrase structure in speech (Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989) as well as music (Krumhansl & Jusczyk, 1990).

The pitch contours of speech seem to have special significance for infants. From at least 1 month of age, infants listen preferentially to infant-directed over adult-directed speech (Cooper, 1990; Cooper & Aslin, 1989; Fernald, 1985; Werker & McLeod, 1989). These types of speech differ substantially in the presence and incidence of *musical* features, that is, features usually associated with music rather than speech. Across a wide range of cultures, infant-directed speech differs from its adult-directed counterpart in a number of features, including its higher pitch, larger pitch range, greater rhythmicity, slower tempo, and distinctive, repeating pitch contours (Fernald & Simon, 1984; Fernald, Taeschner, Dunn, Papoušek, deBoysson-Bardies, & Fukui, 1989; Grieser & Kuhl, 1988; Jacobson, Boersma, Fields, & Olson, 1983; Papoušek & Papoušek, 1981, 1987; Papoušek, Papoušek, & Bornstein, 1985; Stern, Spieker, Barnett, & MacKain, 1983). It is the pitch contours, however, that underlie infants' preference for infant-directed speech (Fernald & Kuhl, 1987).

Infant-directed speech also has notable affective consequences. For example, infants exhibit more positive affect when listening to infant-directed than to adult-directed speech (Werker & McLeod, 1989). Similarly, pitch contours associated with maternal utterances of approval lead to more positive affective displays than those associated with utterances of prohibition (Fernald, in press).

Sensitivity to prosody may be present from the neonatal period (Cooper, 1990), and there are indications that learning in this domain is unusually rapid (see Trehub & Trainor, 1990). Indeed, such skills are likely implicated in infants' differentiation of native from foreign speech (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988) and in recognition of the mother's voice (DeCasper & Fifer, 1980; Mehler, Bertoncini, Barrière, & Jassik-Gerschenfeld, 1978) in the early days of life.

It appears, then, that infants are initially sensitive to suprasegmental or musical aspects of speech, especially pitch contour. Although young infants are also sensitive to segmental aspects of speech (i.e., consonant and vowel sounds; see Kuhl, 1987), effects of experience on phonetic perception are not apparent until 10–12 months of age (Best & McRoberts, 1989; Werker & Lalonde, 1988). At younger ages, infants can differentiate some foreign speech sounds that pose difficulties for their parents (Best & McRoberts, 1989; Burnham, 1986; Trehub, 1976; Werker, 1989; Werker & Lalonde, 1988; Werker & Tees, 1984). This implies that infants' early processing of speech sounds differs from that of adults in having an acoustic or phonetic basis rather than being based on speech sound or phonemic categories. By 1 year of age, there is some perceptual reorganization of speech sound processing such that infants lose this initial "advantage" over adults as they acquire differential sensitivity to some native-language sounds (Best, McRoberts, & Sithole, 1988; Werker & Lalonde, 1988). Presumably, progress toward mature phonemic processing reflects specific linguistic as opposed to general auditory ex-

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the University of Toronto.

Correspondence concerning this article should be addressed to either Laurel J. Trainor or Sandra E. Trehub, Centre for Research in Human Development, University of Toronto, Erindale Campus, Mississauga, Ontario, Canada L5L 1C6.

perience (Best et al., 1988; Werker, in press) and is, therefore, unlikely to be uniform across all speech contrasts. Moreover, there is suggestive evidence that specific cognitive skills emerging at this time play a role in these perceptual developments (Lalonde & Werker, 1990). Thus, linguistic exposure and appropriate cognitive status may be required before certain features of the native language affect perceptual-cognitive structures for speech processing.

There may be parallels in the development of music processing, but little is known about the development of perceptual-cognitive structures relevant to music in general or to the specific musical systems of different cultures. Although features such as pitch contour are essential components of musical structure, they are not unique to music or to any musical system. On the other hand, there are unique sets of rules or conventions that differentiate musical systems, just as there are unique sets of syntactic rules that differentiate languages. For example, the Western tonal system dictates conventional, or prototypic, pitch relations among the notes of a composition as well as the frequency of occurrence of different notes. Moreover, specific combinations of notes generate expectations about the note(s) to follow, with any particular note having a high or low probability of following others. Even adults without formal musical training are sensitive to this structure, being better at processing melodies that exemplify the musical structure of their culture than melodies that do not (Cuddy, Cohen, & Mewhort, 1981; Cuddy, Cohen, & Miller, 1979). For example, Western adults more readily detect changes to a standard melody when that melody conforms to the rules of Western tonality. Of particular interest in the present context is the finding that such changes are more salient when they go outside the original key (i.e., when diatonic structural rules are violated; Cuddy et al., 1979).

Investigators have examined perceptual facilitation, memory, and similarity judgments to reveal adults' knowledge of Western music structure (Bharucha, 1987; Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986, 1987; Cuddy et al., 1979, 1981; Krumhansl, 1979, 1983, 1985, 1990; Krumhansl, Bharucha, & Castellano, 1982; Krumhansl, Bharucha, & Kessler, 1982; Krumhansl & Castellano, 1983; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979). What has emerged is that the processing of musical input seems to be altered by formal or informal exposure to a particular musical system, just as the processing of linguistic input is altered by exposure to a particular language. What is unclear, however, is whether young infants engage in premusical processing of melodies just as they engage in prephonological processing of speech. Prephonological processing leads infants to differentiate native and nonnative contrasts with equal ease, in contrast to later phonological processing, which favors native-language sounds. In analogous fashion, premusical processing could be marked by equivalent processing of musically acceptable and unacceptable changes to melodic patterns, in contrast to musical, or differential, processing of such changes.

It is also the case that certain patterns are inherently easier to process than others. For example, visual patterns conforming to Gestalt principles are easier than nonconforming patterns (Garner, 1974; Younger & Gotlieb, 1988). Similarly, certain melodic patterns are easier to process than others (Boltz & Jones, 1986; Deutsch, 1982; Divenyi & Hirsh, 1974; Jones, 1981; Nickerson & Freeman, 1974). For example, simplicity of contour and hierarchical rule structure have been shown to facilitate processing (Boltz, Marshburn, Jones, & Johnson, 1985). Nevertheless, there is, as yet, no overall theory of pattern "goodness" (see Humphrey & Humphrey, 1989; Trehub & Trainor, 1990, in press). With respect to musical patterns, it is generally acknowledged that pattern "goodness" arises in part from the structure of the auditory system (e.g., primitive grouping processes, see Bregman, 1990; psychoacoustic dissonance, see Terhardt, 1978) and in part from experiential factors (e.g., Dowling & Harwood, 1986; Jones, 1981; Sloboda, 1985), but there is little knowledge about the interaction of primitive and schema-based (i.e., learned) processes.

Cross-cultural research has revealed a few universal features of musical systems (Harwood, 1976) such as the functional equivalence of notes an octave apart (fundamental frequency ratio of 1:2), the greater salience of relational as opposed to absolute information about pitch or duration, and the use of five to seven discrete pitches per octave (Dowling & Harwood, 1986). If music perception follows a developmental course similar to that of speech perception, then young infants should be sensitive to features of musical systems that are universally "good," or easy to learn, whereas older infants or young children should gradually acquire sensitivity to features that are necessarily learned. Thus, a developmental approach to the problem of pattern goodness has the potential to reveal which aspects of musical patterns are intrinsically good and which require more or less exposure to become good.

Western tonal music is based on a division of the octave into 12 discrete pitches separated by equal intervals called semitones (say, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, where Notes 0 and 12 are one octave apart) along a log frequency scale, a division that is by no means universal. Western musical compositions, however, are not based on this equal-interval scale, but rather on an unequal-interval subset. The most common Western scale is the major scale, for example, Notes 0, 2, 4, 5, 7, 9, 11, and 12 (0). It is this specific succession of intervals, rather than absolute pitches, that defines the major scale, so that Notes 3, 5, 7, 8, 10, 12 (0), 14 (2), 15 (3) also form a major scale, as does any interval sequence of 2, 2, 1, 2, 2, 2, 1 semitones. Major scales starting on different notes are called transpositions of each other, as are melodies with the same interval relations between notes. Moreover, different notes of the major scale have different functions. For example, the first note, or tonic, is the most stable, with compositions typically beginning and ending on this note. The fifth note, or dominant, is second in importance. Harmony based on the dominant that is followed or resolved by harmony based on the tonic results in a sense of closure or resolution of tension (Piston, 1969).

There is some evidence that 8- to 10-month-olds show enhanced processing for certain musically well-structured, or good, melodies. For example, Cohen, Thorpe, and Trehub (1987) found that 7- to 11-month-old infants were better at detecting a semitone change upward to the third note of a good, or well-structured, melody based on the major triad (e.g., C E G E C in the key of C major) than a semitone change downward to the third note of a less well-structured, or "bad," melody based on the augmented triad (e.g., C E G♯ E C). These melodies were presented in transposition, precluding the use of absolute pitch cues. Note, moreover, that these two conditions simply reversed which melody was standard and which was the change, or comparison, melody. The finding is of particular interest because the major triad is considered to be a stable, good form in Western music theory, in contrast to the augmented triad, which is considered unstable (i.e., requiring resolution) and dissonant (Piston, 1969). It is not known, however, whether the major triad is an intrinsically good form or good only for those exposed to Western music. In fact, the major triad is not prominent cross-culturally. At the same time, the frequency relations among notes of the major triad (4:5:6) are much simpler than those of the augmented triad (12:15:19), raising the possibility that processing differences arise from the basic structure of the auditory system (Terhardt, 1978).

Goodness of structure was examined further by Trehub, Thorpe, and Trainor (1990), who studied 7- to 10-month-old infants' ability to discriminate (in transposition) a semitone change downward to the fourth note of three different five-note sequences: (a) a good Western melody, in which all notes belonged to a major scale and the implied harmony was typically Western, (b) a bad Western melody, in which the component notes did not belong to any major or minor scale and two intervals were dissonant (tritone), and (c) a non-Western melody, which contained some intervals smaller than a semitone. Although the three standard melodies were similar in their simple up–down contour and pitch range, infants succeeded in detecting the change only in the case of the good Western melody. This suggests that infants' relative ease of processing tone sequences might depend on certain structural features of the patterns. Nevertheless, it is unclear which specific features contribute to pattern goodness for naive infant listeners.

We know, then, that infants process musical aspects of auditory patterns and find some structures better than others, but we know little about the course of development with respect to the specific musical system of exposure. Lynch, Eilers, Oller, and Urbano (1990) compared infants and adults on their detection of a mistuned note in melodies based on the Western major, the Western minor, and the non-Western (Javanese) pélog scales. Although infants did not differ in their detection of mistuned Western and non-Western melodies, adults with limited musical experience and amateur musicians performed better on the Western melodies. Nevertheless, unequivocal interpretation of these findings is precluded for several reasons. First, adults were tested with smaller degrees of mistuning (i.e., smaller pitch differences) than infants, so that direct comparisons of adult and infant performance were not possible. Second, professional musicians might have been expected to show even greater enhancement for Western melodies than less musically experienced adults, but they performed equivalently on Western and non-Western scales. Perhaps professional musicians are simply better at a variety of music-processing tasks. If so, this would imply that extended exposure to a particular system, or musical "talent," not only promotes the development of culture-specific music-processing schemata but also enhances auditory pattern processing in general. Third, Lynch et al. (1990) did not present the standard and comparison melodies in transposition (i.e., different starting pitches), so that discrimination of the single-note change could have been based on absolute as opposed to relative pitch cues. Finally, Lynch et al.'s (1990) conclusion that infants and adults differ in their processing of Western and non-Western scales is based on acceptance of the null hypothesis (i.e., no difference between scale types for infants), which always makes for a weak inference. In any case, the study is important in focusing on potential experiential effects in the processing of musical materials.

The goal of the present research was to examine perceptual processing of Western tonal structure by Western 8-month-old and adult listeners. The standard melody was a well-structured 10-note sequence that conformed to Western tonal conventions. The task of infant and adult listeners was to detect two types of changes to the sixth note of this good melody: (a) a diatonic change (four semitones) that conformed to tonal conventions in the sense that it remained within the key of the melody and the implied dominant harmony at that point, and (b) a nondiatonic change (one semitone) that violated the rules of tonal structure by going outside the key of the melody. Listeners insensitive to tonal structure could be expected to find the four-semitone change (diatonic) more salient than the one-semitone change (nondiatonic). On the other hand, those with tonal knowledge should find the nondiatonic change more salient because of its violation of the preceding tonal context. For such listeners, the diatonic change would represent a subtle variant of the original pattern, one entirely consistent with the underlying structure. Thus, adults were expected to find the nondiatonic change "wrong" or jarring, that is, readily detectable. If 8-month-old infants also found the nondiatonic change more salient than the diatonic change, this would suggest considerable sensitivity to Western tonal structure.

Experiment 1

Method

Subjects. There were 48 healthy, full-term infants between 8 and 9 months of age (27 girls, 21 boys, mean age = 8 months, 11 days). Half of the infants were assigned to the diatonic-change condition and half to the nondiatonic-change condition. An initial group of 28 were given nondiatonic training, and a subsequent group of 20 were given diatonic training. Two additional infants were eliminated for failing to meet the training criterion, and another 5 were eliminated for failing to complete the test session because of fussing.

Apparatus. Testing took place individually in a double-walled, sound-attenuating booth (Industrial Acoustics Co.). The infant was seated on a parent's lap, facing the experimenter, and a loudspeaker and toy box were located 45° to the infant's left. An ECS microcomputer controlled the audio equipment through a custom-made interface. Sine-wave tones were generated by two Hewlett-Packard 3325A synthesizer/function generators and then attenuated by two Med Associates attenuators. The tones were turned on and off, as necessary, by two Med Associates rise/fall switches. Stimuli were presented by means of a Marantz 1070 stereo amplifier and a single Avant 2AX loudspeaker. The experimenter signaled to the computer via a touch-sensitive button box when the infant was ready for a trial (i.e., quiet and looking directly ahead) and when the infant made a 45° (or greater) head turn to the loudspeaker. Four mechanical toys for rewarding correct responses were housed in a four-chamber smoked Plexiglas box under the loudspeaker and were under computer control.

Stimuli. The standard, or background, melody consisted of the tones C_4 E_4 G_4 F_4 D_4 G_3 C_4 E_4 D_4 C_4 in C major (261, 329, 391, 349, 293, 195, 261, 329, 293, and 261 Hz), where the subscripts represent the octave from which a note is drawn (C_4 is middle C; see Figure 1). The tones were sine waves, contiguous, and 400 ms in duration, with 10-ms linear rise and decay ramps. This melody is typically Western, in the sense that all notes belong to one key and the implied harmony is tonic-to-dominant-to-tonic, a common progression in Western music. In both change-type conditions, the contrasting, or change, melody differed from the standard melody only in the middle (sixth) note (see Figure 1). In the diatonic condition, the G_3 of the standard melody (in C major) was changed to B_3 (246 Hz). This is a large change in pitch distance (four semitones) but a small change musically: The B remains within the key of the melody and also within the implied dominant harmony at that point. In the nondiatonic condition, the G_3 (in C major) was changed to $A\flat_3$ (207 Hz). This is a much smaller change in pitch distance (one semitone) but a larger change musically: A♭ goes outside the key of C major.

Infants in each condition were trained with either diatonic or nondiatonic changes. The contrasting, or change, melody in training also differed from the standard melody only in the sixth note, but the change was larger (in pitch distance) than in the test phase. In the case of nondiatonic training, the change note went outside the key of the melody. In the first three training trials, G_3 (in C major) was changed to $C\sharp_3$ (139 Hz), and in all subsequent training trials, it was changed to $E\flat_3$ (156 Hz). In the case of diatonic training, the changed note stayed within the key of the melody. In the first three training trials, G_3 was changed to C_3 (132 Hz), and in all subsequent training trials, to D_3 (147 Hz). Note that the change in pitch distance was greater for diatonic than for nondiatonic training.

The melodies were transposed to three unrelated keys, C major, E major, and A♭ major, and were presented in a quasi-random sequence such that no two consecutive melodies were in the same key. Thus, listeners were forced to use relative pitch information because absolute pitch information was not informative. The intermelody interval was 800 ms. The average intensity level was 75 dB (A), and the ambient noise level was 27 dB (A), or 42 dB (C), measured at the approximate location of the infant's head.

Procedure. A go/no-go conditioned head-turn procedure was used with the infants (see Trehub et al., 1987). The standard melody repeated continuously in transposition (i.e., different starting pitches) from the moment an infant entered the test environment until the infant completed all test trials. The infant's task was to detect any relative pitch (i.e., interval) change in the repeating melody. The experimenter signaled to the computer (via the button box) whenever the infant was quiet and facing directly ahead (i.e., ready for a trial) and also whenever the infant turned 45° or more to the left (toward the speaker). There were two types of trials. Change trials consisted of one presentation of the contrasting, or change, melody in place of the standard melody. Control, or no-change, trials consisted of continued presentation of the standard, or background, melody. A variable number of repetitions of the standard melody separated trials, the minimum being two. Each presentation differed from the preceding presentation in terms of key (i.e., starting note) so that the detection of changes had to be based on relative rather than absolute pitch information. There were 12 control and 12 change trials presented in random order over the course of the test phase. Reinforcement in the form of 4 s of illumination and activation of an animated toy was provided if infants turned on change trials during the response interval that began with the onset of the sixth note of the contrasting melody and ended 3 s later before the onset of the next (i.e., standard) melody. (In comparable test contexts, a response interval of 4–6 s is typical for infant subjects.) The computer recorded head turns during this 3-s response interval on both control and change trials. The experimenter and parent listened to masking music through headphones and were therefore unaware of the type of trial being presented to the infant.

The test phase was preceded by a training phase that was limited to change trials only. Its purpose was to familiarize infants with the procedure—specifically, that head turns to a sound change led to reinforcement. Trials were not signaled in any way (except for the note change on change trials) so that control trials in the test phase were essentially indistinguishable from the background repetitions. During the training phase, failure to respond to a sound change led to the presentation of the next change melody at a level 5 dB greater than the background melody. Correct responding resulted in a 5-dB decrease in the intensity of the next change melody until the background and change melodies were presented at equivalent intensities. Infants were required to make four correct responses in a row with background and change melodies at equivalent intensities before proceeding to the test phase. Testing was terminated for infants who failed to meet this training criterion within 20 trials.

Results and Discussion

For each condition, the proportions of head turns on change trials (hits) and on control, or no-change, trials (false alarms) were transformed to d' scores for each subject according to yes/no tables of signal detection theory (Green & Swets, 1966). This transformation eliminates response bias and ensures that

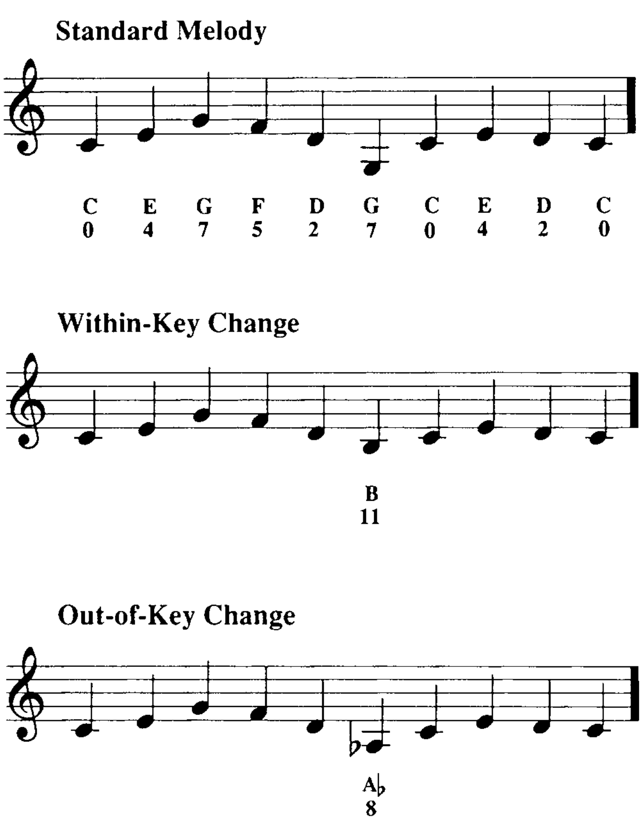

Figure 1. Standard and change melodies in C major. (Note that the diatonic change involves a four-semitone change to the sixth note, whereas the nondiatonic change involves a one-semitone change.)

the distribution of scores meets the assumptions of parametric tests (see Thorpe et al., 1988). Proportions of 0 or 1 can be problematic because they translate into infinite d' scores. The occurrence of infinite d' scores in the current context was presumed to reflect sampling error resulting from limited numbers of trials (12 in this case) rather than statistically infinite d' values (see Macmillan & Kaplan, 1985). To circumvent the problem of infinite d' scores, proportions were calculated by adding ½ to the number of turns (out of 12) and dividing by the number of trials plus 1 (i.e., 13). This adjustment has little effect on d' values and maintains the original ranking of scores (see Thorpe et al., 1988). The maximum d' achievable under these conditions was 3.5.

An analysis of variance (ANOVA) with change type (diatonic vs. nondiatonic) and training type (diatonic vs. nondiatonic) as factors revealed no significant main effects and no significant interaction. Furthermore, infants performed above chance levels in both the diatonic-change condition, $t(23) = 5.60$, $p < .001$ (mean $d' = 0.56$, $SD = 0.49$), and the nondiatonic-change condition, $t(23) = 4.69$, $p < .001$ (mean $d' = 0.46$, $SD = 0.48$), and performance did not differ between these conditions, $t(46) = 0.71$ (see Figure 2, Panel a). Thus, infants did not find the nondiatonic changes easier to detect than the diatonic changes, the trend being in the opposite direction. In addition, type of training had no effect on performance. In the subsequent experiment, we posed the same questions with adults.

Experiment 2

Method

Subjects. There were 28 adults between 18 and 32 years of age (25 women, 3 men) who completed both change-type conditions. Half received diatonic training, and half, nondiatonic training. An additional 4 adults were eliminated for failing to pass training. None of the adults was a professional musician, although some had musical training (16 had 0–4 years of music lessons, 7 had 5–10 years, and 5 had over 10 years).

Apparatus, stimuli, and procedure. The apparatus and stimuli were the same as in Experiment 1. Adults were tested individually in essentially the same procedure as the infants, with as few verbal instructions as possible. As with infants, the standard melody repeated in transposition for the entire test session. In the case of change trials, the comparison melody was substituted for one presentation of the standard melody, after which the standard melody continued. Adults were simply asked to raise a hand (rather than turn) whenever they heard a change in the background melody. Questions were deferred until after the training phase, by which time all adults who met the training criterion understood the task. They were seated facing the experimenter, who listened to music over headphones. Correct responses were rewarded with the animated toys. The procedure differed from that used with the infants, however, in that adults were tested in both change-type conditions. Half of the adults received the diatonic condition first, and half, the nondiatonic condition first. Within each change-type condition, half of the adults received diatonic training, and half, nondiatonic training.

Results and Discussion

An ANOVA on the adult d' scores, with change-type condition (diatonic vs. nondiatonic) as a within-subjects factor and order of conditions (diatonic first vs. nondiatonic first) and training type (diatonic vs. nondiatonic) as between-subjects factors, revealed a significant main effect of change type, $F(1, 24) = 60.92$, $p < .0001$, with nondiatonic changes easier to detect than diatonic changes. Of the 28 adults who completed testing, 13 achieved a perfect score for the nondiatonic change, whereas only 4 did so for the diatonic change. In addition, there were significant interactions between order and training type, $F(1, 24) = 7.18$, $p < .02$, and between order and change type, $F(1, 24) = 17.63$, $p < .003$.

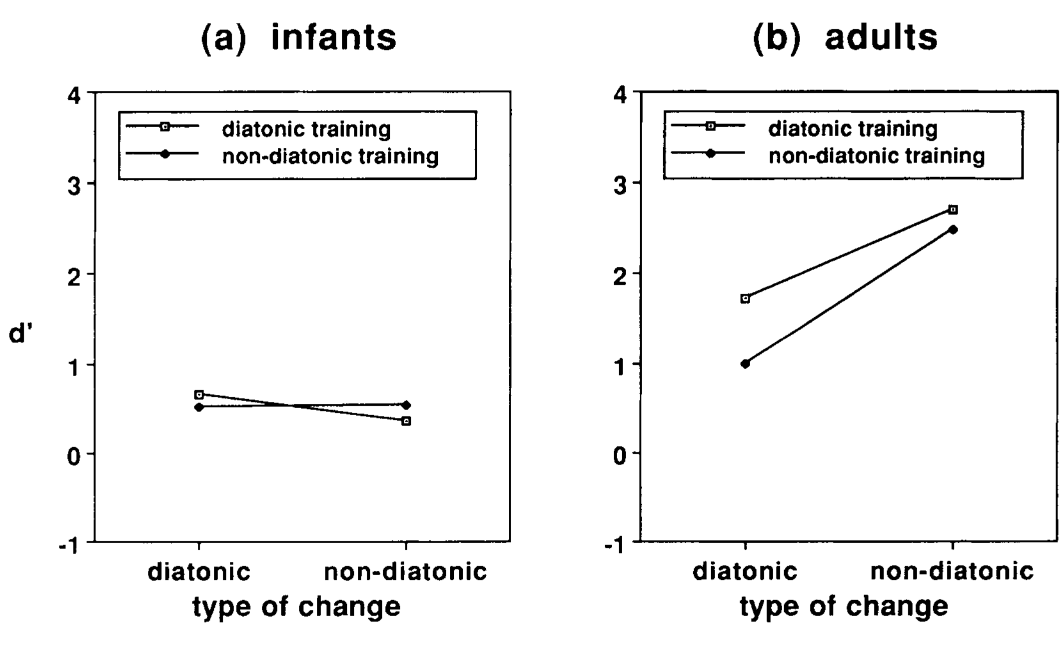

Figure 2. Infant and adult d' scores on the diatonic and nondiatonic change conditions across type of training.

To further investigate these interactions, we performed separate ANOVAS for each training type. The ANOVA for nondiatonic training revealed a significant main effect of change type, $F(1, 12) = 34.17$, $p < .0001$, with nondiatonic changes (mean $d' = 2.48$, $SD = 1.10$) easier to detect than diatonic changes (mean $d' = 0.99$, $SD = 1.54$) (see Figure 2, Panel b). In addition, there was a significant order effect, $F(1, 12) = 10.57$, $p < .01$, with better performance when nondiatonic changes were heard first, and a significant interaction between change type and order, $F(1, 12) = 13.96$, $p < .003$ (see Table 1). Detection of the nondiatonic change was consistently high. By contrast, detection of the diatonic change was at chance levels when this condition was presented first, $t(6) = -1.00$, but well above chance when presented second, $t(6) = 5.53$, $p < .005$. With nondiatonic training, then, adults could perform the diatonic discrimination only after some familiarity with the background melody. However, the ANOVA for diatonic training revealed a significant effect only of change type, $F(1, 12) = 27.17$, $p < .0002$, with nondiatonic changes (mean $d' = 2.71$, $SD = 1.07$) easier to detect than diatonic changes (mean $d' = 1.72$, $SD = 1.17$; see Figure 2, Panel b and Table 1).

Taken together, these results highlight adults' superior detection of nondiatonic changes (i.e., those that violate diatonic structure) over diatonic changes. They indicate, moreover, that detection of the diatonic test changes was enhanced more by training on the difficult-to-detect diatonic changes than by training on the easier nondiatonic changes, $t(12) = 4.75$, $p < .0005$. It is also of interest that 3 of the 4 adults who failed to pass training received diatonic training, even though the pitch changes were greater for diatonic than for nondiatonic training. This implies that diatonic training was indeed more difficult. On the other hand, type of training had no effect on adults' subsequent detection of nondiatonic test changes, $t(12) = 0.70$. This implies that good performance in the diatonic-change condition required extended exposure to the standard pattern (e.g., as in two test sessions) or training on a more difficult task (e.g., the diatonic change), whereas good performance in the nondiatonic-change condition required neither.

Table 1
Adult d' Scores

Order	Type of change: Diatonic	Nondiatonic
Nondiatonic training		
Diatonic first		
M	−0.25	2.19
SD	0.67	1.10
Nondiatonic first		
M	2.23	2.76
SD	1.07	1.10
Diatonic training		
Diatonic first		
M	1.75	3.13
SD	0.89	0.49
Nondiatonic first		
M	1.68	2.30
SD	1.47	1.36

Although no participant was an active amateur or professional musician, the participants did differ in musical training. Simple regressions revealed no significant relations between musical training (i.e., number of years of music lessons) and relative performance (difference between diatonic and nondiatonic conditions) in either the nondiatonic-training condition or the diatonic-training condition. Thus, formal musical training was unnecessary for superior detection of nondiatonic over diatonic changes.

On the basis of relative motivation and attention, one would expect adults to perform better than infants on almost any experimental task. In fact, this was the case overall. However, infants actually performed better than adults in one condition. Recall that infants were tested in only one change-type condition, whereas adults were tested in both. Because adult performance improved in some conditions between the first and second test session, it is appropriate to restrict adult–infant comparisons to data obtained from the first test session (the only session for infants). Comparisons of performance on the detection of diatonic changes in the context of nondiatonic training revealed significant differences favoring infants, $t(19) = 2.78$, $p < .01$, differences that are highly suggestive of differential processing. In short, adults more readily detected a melodic change that violated diatonic structure than one that did not, but infants' performance was unrelated to tonal structure.

General Discussion

Adults easily detected the change that violated Western musical structure (i.e., nondiatonic) but had difficulty with the change that preserved such structure (i.e., diatonic). Infants, on the other hand, detected both changes equally well. Moreover, only adults were affected by the type of training (diatonic vs. nondiatonic). Specifically, adult detection of the diatonic change was better when preceded by diatonic training than by nondiatonic training. By contrast, their detection of nondiatonic changes was consistently good and unrelated to the nature of training. At issue here is not whether adults are capable of acquiring the diatonic discrimination. It is likely that sufficient exposure to the standard melody or practice on change trials would lead to excellent performance. The central issue, rather, is one of *response proclivities* as opposed to *sensory capabilities* (Espinoza-Varas & Watson, 1989, pp. 85–88). Adults simply found the diatonic discrimination more difficult than the nondiatonic.

Adults' difficulty with the diatonic discrimination is especially noteworthy in the context of a task designed for infants. In the typical melody discrimination task (e.g., Bartlett & Dowling, 1980), adults are presented with one instance of a standard melody and one instance of a comparison melody (in transposition) and they are required to judge whether the second is an exact transposition of the first (i.e., the same melody). On subsequent trials, different exemplars are used as standard and comparison melodies to prevent adults from becoming familiar with the standard melody (i.e., learning the pattern of intervals). In the present study, however, adults heard many repetitions of the standard melody during the training phase and dozens more by the end of the first test session. Indeed, variants of this procedure, with its numerous repetitions of the standard pattern and interspersed comparisons, have made it possible for adults to perform within-category phoneme discriminations that are difficult or impossible with conventional psychophysical procedures (e.g., Carney, Widin, & Viemeister, 1977). In this context, infants' superiority over adults in detecting the diatonic change (with nondiatonic training) provides compelling evidence of melodic processing differences in infancy and adulthood.

It is likely that adults' perception of melodies is influenced by schemata, or perceptual-cognitive structures, specific to Western tonal structure. Such schemata would increase the difficulty of detecting changes that are consistent with diatonic conventions and the implied harmony of the melody. Without hearing the melodies of the study, it is difficult to appreciate the distinct perceptual effects arising from the diatonic and nondiatonic changes. Adult nonparticipants, who were simply presented with one or two examples of each type of change (with the locus of change identified by the experimenter), confidently asserted (wrongly, of course) that the diatonic change was substantially smaller (in pitch distance) than the nondiatonic. Infants, on the other hand, were seemingly unaffected by the conformance or nonconformance of the change to Western diatonic conventions. In previous research, infants have more readily detected a change to some melodies than to others (Cohen et al., 1987; Trehub et al., 1990). Of interest is the fact that the melodies associated with enhanced infant processing conformed to ideals of Western diatonic structure, raising the possibility that Western musical structure could be acquired with limited musical exposure. An alternative explanation is that the prototypically Western melodies in these studies embodied features that made them intrinsically easy to process independent of their "Westernness." One potentially relevant feature is the simple frequency ratios of important intervals, a feature that characterized melodies associated with better infant performance but not those associated with discrimination failures. According to this explanation, the simple frequency ratios or some other feature of the good melodies enhanced encoding. The bad melodies in these studies were similar in contour and frequency range but did not possess the critical good features and were not well encoded.

The results of the present study confirm earlier findings but also extend such findings by adding credence to the notion that some melodies are intrinsically good. The confirmation comes from infants' ability to detect interval changes in the context of a good, or well-structured, melody that was considerably more complex than the five-tone good melodies in previous studies. If infants' ability to detect such changes was dependent on musical exposure, then such exposure should have led to superior detection of the change that violated Western musical structure over the one that conserved such structure. The finding of equivalent performance on both types of changes suggests that the nondiatonic change is not intrinsically more discrepant than the diatonic but becomes so on the basis of exposure to Western music (as with adults). If exposure is not relevant to infants' performance in the present investigation, it is unlikely to be relevant to infants' enhanced processing of well-structured melodies in previous studies.

Although young infants are sensitive to a number of features common to speech and music, some schemata, or perceptual-cognitive structures, specific to the language (e.g., Werker & Lalonde, 1988) or musical system of their culture do not emerge until later. This raises questions about the importance of musical processing in early life and the possible relations between music and language development. In some ways, infants' and adults' processing of auditory patterns is remarkably similar. For example, infants' sensitivity to relational (i.e., contour) information as opposed to absolute pitch information is similar to adult processing strategies and very different from the pitch processing strategies of various nonhuman species such as songbirds (Hulse, Page, & Braaten, 1990) and monkeys (D'Amato, 1988), which depend on absolute as opposed to relational cues. Fernald (1989, in press) has suggested that the pitch contours of adult speech to infants represent the infant's earliest associations between sound patterns and meaning, with similar contours evident across disparate languages (Fernald et al., 1989; Grieser & Kuhl, 1988). Moreover, distinctive contours are seen in specific caregiving situations—for example, falling contours to soothe infants and rising, or bell-shaped, contours to capture their attention (Papoušek & Papoušek, 1981; Stern, Spieker, & MacKain, 1982).

Perhaps music and speech are intimately connected in early life, with common perceptual processing mechanisms for musical features. In light of the clearly demonstrated salience of pitch contour in infant-directed speech, Fernald (1989) posed an intriguing question: "Is the melody the message?" In considering this question, it is important to note that pitch contour alone does not constitute melody. Rather, melodies exist within the more complex structure of unique musical systems that define a small number of pitch categories, their relative importance, and the syntactic and semantic relations among them. Theoretically, music and language could interact developmentally in several ways. From the beginning, they might be independent systems, such that once the input is identified as speech or music, processing would continue within the appropriate system. Alternatively, a single system for processing musical features might be operative until specialized music and speech processors emerge during the 1st year of life or beyond. Further research on the development of culture-specific schemata for speech and music should enhance our understanding of auditory pattern processing and its development.

References

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 501–515.

Best, C. T., & McRoberts, G. W. (1989, April). *Phonological influence on infants' perception of two nonnative speech contrasts.* Paper presented at the meeting of the Society for Research in Child Development, Kansas City, MO.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal* *of Experimental Psychology: Human Perception and Performance, 14,* 45–60.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5,* 1–30.

Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13,* 63–102.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 403–410.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41,* 519–524.

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18,* 389–431.

Boltz, M., Marshburn, E., Jones, M. R., & Johnson, W. W. (1985). Serial-pattern structure and temporal-order recognition. *Perception & Psychophysics, 37,* 209–217.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics, 7,* 207–240.

Carney, A. E., Widin, G. P., & Viemeister, N. F. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America, 62,* 961–970.

Chang, H. W., & Trehub, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology, 24,* 324–331.

Cohen, A. J., Thorpe, L. A., & Trehub, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology, 41,* 33–47.

Cooper, R. B. (1990, April). *Preference for infant-directed speech in newborn infants.* Paper presented at the International Conference on Infant Studies, Montreal, Quebec, Canada.

Cooper, R. B., & Aslin, R. N. (1989). The language environment of the young infant: Implications for early perceptual development. *Canadian Journal of Psychology, 43,* 247–265.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 869–883.

Cuddy, L. L., Cohen, A. J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology, 33,* 148–156.

D'Amato, M. R. (1988). A search for tonal pattern perception in cebus monkeys: Why monkeys can't hum a tune. *Music Perception, 5,* 453–480.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208,* 1174–1176.

Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development, 5,* 261–276.

Deutsch, D. (1982). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (pp. 271–316). New York: Academic Press.

Divenyi, P. L., & Hirsh, I. J. (1974). Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America, 56,* 144–151.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* Orlando, FL: Academic Press.

Espinoza-Varas, B., & Watson, C. S. (1989). Perception of complex auditory patterns by humans. In R. J. Dooling & S. H. Hulse (Eds.), *The comparative psychology of audition: Perceiving complex sounds* (pp. 67–94). Hillsdale, NJ: Erlbaum.

Ferland, M. B., & Mendelson, M. J. (1989). Infants' categorization of melodic contour. *Infant Behavior and Development, 12,* 341–355.

Fernald, A. (1985). Four-month-old infants prefer to listen to Motherese. *Infant Behavior and Development, 8,* 181–195.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60,* 1497–1510.

Fernald, A. (in press). Meaningful melodies in mothers' speech to infants. In H. Papoušek, V. Jurgens, & M. Papoušek (Eds.), *Origins and development of nonverbal vocal communication: Evolutionary, comparative, and methodological aspects.* New York: Cambridge University Press.

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese. *Infant Behavior and Development, 10,* 279–293.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113.

Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16,* 477–501.

Garner, W. R. (1974). *The processing of information and structure.* Hillsdale, NJ: Erlbaum.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24,* 14–20.

Harwood, D. L. (1976). Universals in music: A perspective from cognitive psychology. *Ethnomusicology, 20,* 521–534.

Hulse, S. H., Page, S. C., & Braaten, R. F. (1990). An integrative approach to serial pattern learning: Music perception and comparative acoustic perception. In W. C. Stebbins & M. Berkley (Eds.), *Comparative perception, Vol. 2: Communication* (pp. 3–34). New York: Wiley.

Humphrey, G. K., & Humphrey, D. (1989). The role of structure in infant visual pattern perception. *Canadian Journal of Psychology, 43,* 165–182.

Jacobson, J. L., Boersma, D. C., Fields, R. B., & Olson, K. L. (1983). Paralinguistic features of speech to infants and small children. *Child Development, 54,* 436–442.

Jones, M. R. (1981). Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology, 1,* 14–31.

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16,* 66–68.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11,* 346–374.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception, 1,* 28–62.

Krumhansl, C. L. (1985). Perceiving tonal structure in music. *American Scientist, 73,* 371–378.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C. L., Bharucha, J. J., & Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics, 32,* 96–108.

Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three related keys. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 24–36.

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition, 11,* 325–334.

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science, 1,* 70–73.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334–368.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–594.

Kuhl, P. K. (1987). Perception of speech and sound in early infancy. In P. Salapatek & L. Cohen (Eds.), *Handbook of infant perception, Vol. 2: From perception to cognition* (pp. 275–382). Orlando, FL: Academic Press.

Lalonde, C., & Werker, J. (1990, April). *Cognitive/perceptual integration of three skills at 9 months.* Paper presented at the International Conference on Infant Studies, Montreal, Quebec, Canada.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1,* 272–276.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98,* 185–199.

Mehler, J., Bertoncini, J., Barrière, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception, 7,* 491–497.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29,* 143–178.

Nickerson, R. S., & Freeman, B. (1974). Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation and extensive practice. *Perception & Psychophysics, 16,* 471–477.

Papoušek, H., & Papoušek, M. (1987). Intuitive parenting: A dialectic counterpart to the infant's integrative competence. In J. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 669–720). New York: Wiley.

Papoušek, M., & Papoušek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. In L. P. Lipsitt (Ed.), *Advances in infancy research* (Vol. 1, pp. 163–224). Norwood, NJ: Ablex.

Papoušek, M., Papoušek, H., & Bornstein, M. H. (1985). The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 269–297). Norwood, NJ: Ablex.

Piston, W. (1969). *Harmony.* New York: Norton.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music.* Oxford, England: Clarendon Press.

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language, 10,* 1–15.

Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology, 18,* 727–735.

Terhardt, E. (1978). Psychoacoustical evaluation of musical sounds. *Perception & Psychophysics, 23,* 483–492.

Thorpe, L. A. (1986, April). *Infants categorize rising and falling pitch.* Paper presented at the International Conference on Infant Studies, Los Angeles, CA.

Thorpe, L. A., & Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology, 25,* 122–127.

Thorpe, L. A., Trehub, S. E., Morrongiello, B. A., & Bull, D. (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology, 24,* 484–491.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47,* 466–472.

Trehub, S. E. (1987). Infants' perception of musical patterns. *Perception & Psychophysics, 41,* 635–641.

Trehub, S. E. (1989). Infants' perception of musical sequences: Implications for language acquisition. *Journal of Speech-Language Pathology and Audiology, 13,* 3–11.

Trehub, S. E. (1990). The perception of musical patterns by human infants: The provision of similar patterns by their parents. In M. A. Berkley & W. C. Stebbins (Eds.), *Comparative perception, Vol. 1: Basic mechanisms* (pp. 429–459). New York: Wiley.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development, 55,* 821–830.

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1985). Infants' perception of melodies: Changes in a single tone. *Infant Behavior and Development, 8,* 213–223.

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1987). Organizational processes in infants' perception of auditory patterns. *Child Development, 58,* 741–749.

Trehub, S. E., Thorpe, L. A., & Trainor, L. J. (1990). Infants' perception of *good* and *bad* melodies. *Psychomusicology, 9,* 5–15.

Trehub, S. E., & Trainor, L. J. (1990). Rules for listening in infancy. In J. Enns (Ed.), *The development of attention: Research and theory* (pp. 87–119). Amsterdam: Elsevier.

Trehub, S. E., & Trainor, L. J. (in press). Listening strategies in infancy: The roots of language and musical development. In S. McAdams & E. Bigand (Eds.), *Cognitive aspects of human audition.* Cambridge, England: Cambridge University Press.

Werker, J. F. (1989). Becoming a native listener. *American Scientist, 77,* 54–59.

Werker, J. F. (in press). The ontogeny of speech perception. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception.* Hillsdale, NJ: Erlbaum.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24,* 672–683.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43,* 230–246.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49–63.

Younger, B., & Gotlieb, S. (1988). Development of categorization skills: Changes in the nature or structure of infant form categories. *Developmental Psychology, 24,* 611–619.

Received June 19, 1990
Revision received May 15, 1991
Accepted May 17, 1991 ■