MUSIC AND SPEECH PROCESSING IN THE FIRST YEAR OF LIFE

Sandra E. Trehub

CENTRE FOR RESEARCH IN HUMAN DEVELOPMENT
ERINDALE COLLEGE
UNIVERSITY OF TORONTO
MISSISSAUGA, ONTARIO, CANADA L5L 1C6

Laurel J. Trainor

DEPARTMENT OF PSYCHOLOGY
MCMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA L8S 4K1

Anna M. Unyk

CENTRE FOR RESEARCH IN HUMAN DEVELOPMENT
ERINDALE COLLEGE
UNIVERSITY OF TORONTO
MISSISSAUGA, ONTARIO, CANADA L5L 1C6

I. WHY STUDY MUSIC AND SPEECH PROCESSING IN INFANCY?
  A. SPEECH AND EMOTION
  B. MUSIC AND EMOTION
  C. FOCUS OF THE ESSAY

II. SPEECH TO INFANTS
  A. FEATURES OF INFANT-DIRECTED SPEECH
  B. SITUATIONAL, INDIVIDUAL, AND CROSS-CULTURAL DIFFERENCES
  C. THE LISTENER'S CONTRIBUTION
  D. INFANT-DIRECTED SPEECH AS COMMUNICATION

III. SONG TO INFANTS
  A. LULLABIES: DESCRIPTIVE RESEARCH ACROSS CULTURES
  B. LULLABY IDENTIFICATION ACROSS CULTURES
  C. MATERNAL SINGING: THE INFANT'S CONTRIBUTION
  D. PATERNAL SINGING

IV. INFANTS' PROCESSING OF SPEECH SEQUENCES
  A. ATTENTIONAL PREFERENCES
  B. AFFECTIVE PREFERENCES
  C. TEMPORAL PATTERNING AND ATTENTIONAL PREFERENCES

ADVANCES IN CHILD DEVELOPMENT
AND BEHAVIOR, VOL. 24

V. INFANTS' PROCESSING OF MUSICAL SEQUENCES
A. ATTENTIONAL PREFERENCES
B. METHODOLOGICAL STRATEGY
C. PITCH PATTERNING
D. TEMPORAL PATTERNING
E. MELODY AND CONTOUR
F. "GOOD" AND "BAD" MELODIES
G. GENERAL PERSPECTIVES

VI. MUSIC AND SPEECH: PRESENT PERSPECTIVES AND FUTURE PROSPECTS
A. CAREGIVERS' SPEECH AND SONG
B. INFANTS' SPEECH AND SONG PROCESSING

REFERENCES

I. Why Study Music and Speech Processing in Infancy?

Little justification is required for the study of speech processing in infancy. After all, mothers and other caregivers generally talk to noncomprehending infants, and ultimately, these infants become children with adult-like speech reception and production skills. In this light, what could be more reasonable than to explore very early abilities, processing predispositions, and the course of early learning in this domain?

The case for the study of music processing in infancy seems much less compelling, largely because of our ignorance about the prevalence and functions of music. Music, like language, has been found in every culture and historical period. Moreover, music has typically occupied a central role in work and play, involving all members of society (Bebey, 1969). Indeed, the Western inclination to relegate music making to talented performers and isolated artistic contexts (e.g., concerts) is unusual when viewed in historical and cross-cultural perspective (Walker, 1990). Not only is participation in music widespread, but so is belief in its power over physical and mental states (Walker, 1990). In this light, posing questions about early processing skills, predispositions, and learning in relation to music seems to be as reasonable as posing such questions in relation to speech.

Speech and music processing may be linked in infancy in ways that go beyond the obvious sharing of reception and production channels (i.e., ears and voice). For prelinguistic infants, speech is linguistic only from the perspective of the talker or mature observer. From the very young infant's perspective, speech consists of complex patterns of semantically meaningless sounds that vary over time. To the extent that such patterns are meaningful in the earliest months of

life, these meanings are necessarily affective or emotional (Fernald, 1989, 1992; Lewis, 1951).

A. SPEECH AND EMOTION

Vocal but nonverbal expressions of emotion may be universal (for a review see Frick, 1985). Prosodic contours in speech (i.e., patterns of pitch and loudness over time) are thought to reflect the talker's emotional state (Knower, 1941; Williams & Stevens, 1972). For example, falling contours are correlated with pleasantness and rising contours with surprise or fear (Scherer, 1974). Spectral structure (i.e., amount of energy at different frequencies), which confers a distinctive sound quality, tone of voice, or timbre, is also important for emotion recognition (Lieberman & Michaels, 1962). For example, the emotions of pleasantness and happiness are associated with relatively few component frequencies (Scherer & Oshinsky, 1977). A speaking voice with few component frequencies sounds "pure," more like a flute than a piano, for example. Also, facial expressions such as smiling affect the shape of the vocal tract and the resultant voice quality (Laver, 1980). Positive emotions tend to shift energy to higher component frequencies relative to negative emotions (Frick, 1985). Moreover, as a talker's pitch level is raised relative to his or her usual level, vowel spectral noise may decrease, leading to a corresponding decrease in listeners' perception of vowel roughness (Emanuel & Smith, 1974; Newman & Emanuel, 1991). Presumably, this higher-pitched voice would have a more melodious quality. Pitch level also contributes to perceived emotion. High pitch signals happiness and friendliness for human interactants and nonaggression or submission for human and nonhuman interactants (Frick, 1985; Stross, 1977).

Prosody, then, may be the principal carrier of emotion in speech, one that is not primarily dependent on symbolic understanding or cultural conventions (Frick, 1985). As a result, listeners can identify the emotional tone of content-filtered speech (Scherer, Koivumaki, & Rosenthal, 1972) and can derive similar interpretations of nonverbal emotional content across cultures (Krauss, Curran, & Ferleger, 1983).

B. MUSIC AND EMOTION

The view of music as the universal language of the emotions (Langer, 1957) is an oversimplification because music also engages culture-specific cognitive structures. Nevertheless, music can express or represent emotions (Langer, 1957; Meyer, 1956; Trainor & Trehub, 1992b) and can also alter the emotional state of the mature listener (Francès, 1958; Meyer, 1956). Emotion, then, could be regarded as a quality that permeates music and speech. From the perspective of "premusical" or musically unacculturated infants, music, like speech, consists

of complex sound patterns that vary over time. For infants, moreover, musical patterns must bear considerable similarity to the prosody of speech, notably its intonation, rhythm, and stress (Trehub, 1990). For prelinguistic and premusical listeners, then, patterns of speech and music might well engage common processing strategies.

C. FOCUS OF THE ESSAY

In the present article, we focus on potential similarities between speech and music from the perspective of infant listeners. The stimuli of concern are sound sequences rather than single sounds, despite the predominant research focus on the latter class of stimuli. The exclusion of single sounds can be justified on a number of grounds. First, the literature contains several comprehensive reviews of infants' ability to perceive single speech and nonspeech sounds (e.g., Aslin, Pisoni, & Jusczyk, 1983; Jusczyk, 1992; Kuhl, 1988; Schneider & Trehub, 1992; Werker, 1991). Second, and perhaps more important, evidence indicates that global patterns of speech are more salient in the prelinguistic period than are individual speech segments (i.e., consonants and vowels) (Crystal, 1973; Lewis, 1951). In the nonspeech domain, evidence also indicates that infants proceed from global processing of auditory patterns to local processing of pattern details (Morrongiello, 1988; Trehub, 1985, 1990; Trehub & Trainor, 1990).

In drawing parallels between speech and music, we focus on two principal issues: the input provided by caregivers for their infants and the processing of such input by infant listeners. Much of the work to be reported, particularly in the musical domain, is relatively recent. As a result, the exposition is tentative rather than definitive, its purpose being to suggest new avenues for future research and thinking.

II. Speech to Infants

Infant-directed speech, often termed *motherese* or *baby talk,* tends to be elicited in caregivers by the mere presence of an awake infant (Fernald, 1984; Fernald & Simon, 1984; Rheingold & Adams, 1980; Rosenthal, 1982). Although the principal focus has been on its syntactic structure and suitability as a medium for language teaching (Murray, Johnson, & Peters, 1990; Newport, Gleitman, & Gleitman, 1977; Sherrod, Friedman, Crawley, Drake, & Devieux, 1977; Snow, 1977), researchers have shown growing interest in its prosodic form (e.g., pitch, intonation, stress, rhythm) and suitablilty for attention and affect regulation (Fernald, 1984, 1985, 1989; M. Papoušek, Papoušek, & Symmes, 1991; Stern, Spieker, Barnett, & MacKain, 1983). These suprasegmental or prosodic aspects

are the variables that have the greatest potential relevance to prelinguistic listeners.

A. FEATURES OF INFANT-DIRECTED SPEECH

The differences between infant-directed and adult-directed speech are considered to exceed those between different dialects of a language. The most notable features of this special speech register for infants are its increased pitch, greater pitch range, elongated vowels, simpler pitch contours, rhythmic regularity, slower tempo, briefer utterances, and overall repetitiveness compared with standard adult speech (Beebe, Feldstein, Jaffe, Mays, & Alson, 1985; Ferguson, 1964; Fernald & Simon, 1984; M. Papoušek, Papoušek, & Bornstein, 1985; Sachs, 1977; Stern et al., 1983; Stern, Spieker, & MacKain, 1982). The average pitch of the caregiver's voice rises by about three or four semitones (25–33%) in interactions with infants (Fernald & Simon, 1984; Jacobson, Boersma, Fields, & Olson, 1983; Papoušek et al., 1985), generating a falsetto quality. The most celebrated aspect of infant-directed speech is the unique set of contours that apparently characterize such communications. In contrast to the narrow pitch range and multiple directional changes of pitch movement that characterize adult-directed speech (Bolinger, 1970), very simple contours with an expanded pitch range are typical of infant-directed speech (Fernald & Simon, 1984). These contours are usually unidirectional (rising or falling in pitch) but sometimes bidirectional (rise–fall or fall–rise) and they are used over and over with different component words, phrases, and sounds (H. Papoušek & Papoušek, 1984). Poor articulation is often evident and utterances may consist solely of elongated vowels or consonant–vowel syllables stretched out over one of these expanded contours (Fernald & Simon, 1984; M. Papoušek & Papoušek, 1981). Some idealized examples are provided in Fig. 1.

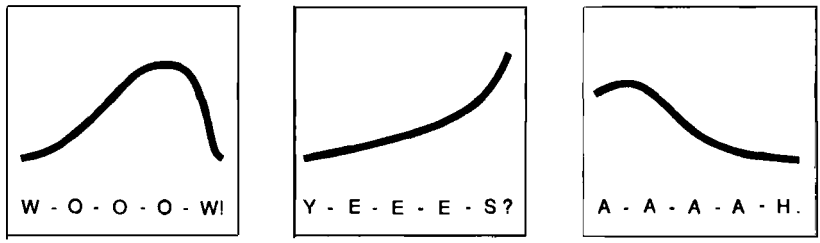

Fig. 1. Schematic illustrations of hypothetical infant-directed utterances. Wow *is greatly extended over a rise–fall contour,* yes *over a rising contour, and* ah *over a falling contour. These contours likely describe correlated changes in loudness and pitch.*

B. SITUATIONAL, INDIVIDUAL, AND CROSS-CULTURAL DIFFERENCES

These pitch contours seem to be roughly tied to specific situational contexts. For example, rising contours tend to be used to capture an infant's attention and to elicit some response (Ferrier, 1985; M. Papoušek et al., 1991). Bell-shaped contours are sometimes used for attention capture as well, but they are used more frequently for maintaining such attention and expressing approval (Fernald, 1989; Stern et al., 1982). By contrast, sustained or falling contours are used for soothing infants and promoting sleep (M. Papoušek & Papoušek, 1981; M. Papoušek et al., 1991). Rhythm and tempo are also tied to particular contexts, with increasing tempo for inattentive infants, decreasing tempo for infants progressing toward sleep, and varying rhythm and tempo for fussy infants (M. Papoušek & Papoušek, 1981).

If speech adjustments of this nature were limited to North American mothers, they would be indicative of cultural variations in caregiving, but they would be of less general interest; however, similar modifications can be seen in a wide range of structurally different languages including German, Arabic, Mandarin Chinese, Japanese, Spanish, Kwar'ae, and British English (Ferguson, 1964; Fernald et al., 1989; Grieser & Kuhl, 1988; Papoušek et al., 1991; Watson-Gegeo & Gegeo, 1986). Moreover, fathers, primiparous mothers, male and female nonparents, and even preschool children with and without younger siblings adjust their speech in relation to infant listeners (Anderson, 1986; Dunn & Kendrick, 1982; Fernald & Simon, 1984; Jacobson et al., 1983; Rheingold & Adams, 1980; Sachs & Devin, 1976; Watson-Gegeo & Gegeo, 1986; Weeks, 1971), although the nature and extent of such adjustments vary somewhat. For example, fathers alter their vocal behavior less than do mothers, omitting the characteristic increase in pitch range (Fernald et al., 1989). On the whole, however, speech adjustments in relation to infants appear to be independent of language, culture, and caregiving experience.

This conclusion is somewhat of an overstatement, however. Notable cross-cultural variations have been found and they may reflect differences in expressiveness, on the one hand, and caregiving attitudes, on the other. For example, Fernald et al. (1989) described more extreme intonational modifications in American English compared with British English, French, Italian, German, and Japanese speech to infants. Other comparisons of American English with British English (Shute & Wheldall, 1989) and Mandarin Chinese (Grieser & Kuhl, 1988; M. Papoušek et al., 1991) are in line with this finding. Fernald (1992) suggested that such differences reflect variations in cultural display rules, with Asian mothers in particular considering vocal and facial expressiveness less socially acceptable than do European and American mothers. Other research indicates quite the

opposite, that Japanese mothers may be more affect oriented and American mothers more information oriented (Toda, Fogel, & Kawai, 1990). Indeed, Japanese mothers use repetitive nonsense syllables, onomatopoeia, invented lexical items, and imitations of their infant's sounds more than do American mothers (Morikawa, Shand, & Kosawa, 1988; Toda et al., 1990), which lends credence to the latter view. Expanded intonation contours may not be the only effective means of recruiting infants' attention despite their wide usage by North American mothers. Moreover, Japanese mothers seem to spend more time soothing their infants (by vocalization and touch) and less time arousing them compared with American mothers, which may reflect different cultural values and caretaking goals (Caudill & Weinstein, 1969; Toda et al., 1990). More soothing than arousing is also found in the Gusii tribe in the highlands of Kenya (Dixon, Tronick, Keefer, & Brazelton, 1981). Similarly, the predominance of falling over rising contours in Mandarin Chinese (Grieser & Kuhl, 1988) is consistent with the notion of soothing having priority over arousing in the care of infants.

Because most research on infant-directed speech has been focused on highly industrialized societies, North America and Western Europe in particular, the view of typical infant-directed speech as arousing and attention getting may be unwarranted. A question that remains is whether arousing or soothing vocalizations predominate in more broadly based samples of caregivers, particularly those in "traditional" societies, in which mothers and infants spend most of their waking and sleeping hours together.

C. THE LISTENER'S CONTRIBUTION

The prevailing wisdom is that speech modifications to infants are effected more or less intuitively (M. Papoušek et al., 1985). The point is not that adults totally lack awareness of such behavior but rather that at least some aspects of this behavior are not manipulated consciously. As a result, situations of simulated as opposed to actual interactions with infants tend to induce less than the full set of infant- or child-directed adjustments (Fernald & Simon, 1984; Jacobson et al., 1983; Murray & Trevarthan, 1986). Why might this difference arise? The behavior of mothers (and other caregivers) may be driven, at least in part, by feedback from the infant. On the one hand, fluctuations in the infant's attention and arousal provide cues to the fine-tuning of maternal behavior. On the other hand, the infant's presence and behavior may trigger alterations in the mother's own state that are reflected in her behavior. It is likely that both factors are at work, although the latter may generate the most noticeable differences in vocal quality. In this regard, one can think of the speech of lovers, which often exhibits infant- or child-directed qualities such as heightened pitch, the use of dimin-

utives, and nonsense sounds. Intonational stereotypy can also be seen in some adult-directed utterances such as greetings (e.g., *Hi* stretched out over a bell-shaped contour) and exclamations (*Wow!*).

Infant-directed speech is tailored considerably to the age and ability of the listener. For example, the full complement of infant-directed modifications seems to reach its height when the listener is about 4 months of age, becoming somewhat attenuated thereafter (Stern et al., 1983). Many adjustments, however, remain in evidence well beyond infancy (Ferguson, 1964; Garnica, 1977; Snow, 1977), with different parameters undergoing selective enhancement or attenuation. For example, variations in articulatory distinctiveness (Malsheen, 1980), emphatic stress (Fernald & Mazzie, 1991), and the use of diminutives (Jocić, 1978) are linked to the child's language reception and production skills.

D. INFANT-DIRECTED SPEECH AS COMMUNICATION

In line with the notion of maternal prosody as emotionally expressive (Lewis, 1951), Fernald (1989) explored adult listeners' ability to discern different intentions from such utterances. She tape-recorded mothers as they interacted with their 10- to 14-month-old infants in play and caretaking contexts. Five broad communicative categories or utterance types were selected for study: (1) approving, in response to the infant's completion of a simple task; (2) prohibiting, designed to prevent the infant from touching an object about to be touched; (3) comforting an upset infant; (4) attention bidding, which involved directing the infant's attention to an object or action; and (5) game initiating, which involved a hiding game such as *peek-a-boo*. She also tape-recorded the same women as they interacted in somewhat comparable contexts (e.g., soothing, warning) with their husbands. These utterances were content filtered (i.e., frequencies above 400 Hz removed) so that they became semantically unintelligible while still retaining their overall prosodic form. As Fernald (1989) noted, they sounded like "speech heard through a wall" (p. 1503).

Adult listeners were required to assign each utterance to one of the five aforementioned communicative categories and seemed to do so with relative ease. Performance, which was well above chance for each category, was unrelated to listeners' parental status, age, and experience with children. Although adults could also discern similar communicative intent from nonmaternal (adult-directed) prosody, they were considerably less accurate for all categories except comforting vocalizations. From the perspective of adult listeners, then, infant-directed prosody is for the most part more informative than adult-directed prosody. The greater transparency of infant-directed prosody may be linked to its greater simplicity or to its decoupling from conventional communicative symbols.

III. Song to Infants

Given the rich descriptions of maternal vocal behavior and the characterization of much of this behavior as musical or melodic (Fernald, 1989, 1992; M. Papoušek & Papoušek, 1981), we were surprised that interest in this domain had not extended to maternal song. The practice of singing to infants appears to be widespread, perhaps universal. In fact, one ubiquitous musical genre, the lullaby, is regularly sung to soothe infants and induce sleep (Brakeley, 1950; Cass-Beggs & Cass-Beggs, 1969). Detailed descriptive material on lullabies is sparse but the available research provides suggestive evidence of considerable similarity between infant-directed speech and song.

A. LULLABIES: DESCRIPTIVE RESEARCH ACROSS CULTURES

McCosker (1974) has provided one of the more extensive descriptions of lullaby form and function in his work on the Cuna Indians of Panama. The singer of Cuna lullabies enjoys greater freedom for improvisation of text and melody compared with other songs. Unlike other songs, lullabies are of indefinite length, continuing until their function has been fulfilled. Cuna lullaby lyrics reveal considerable word reduplication, sequence repetition, and common words incorporated into repetitive rhythmic patterns, as illustrated (below) by a lullaby from Nalunega Island (McCosker, 1974, pp. 126–127). Translations by a native Cuna speaker are provided (see McCosker, 1974, p. 146). In some cases, the translator had difficulty with words, and rough glosses are provided in parentheses.

(1) a. *pani kala pani poa nai tai ye*
little ones come here to the hammock
b. *poe pii poe pii pani tai maloye*
you are crying come here
c. *nana peka u kachi pa kine*
mama is always here in the hammock
d. *nana peka nai kucha pani nukku pa kine pani poa tii pa ye um*
mama sees you crying and will take you in her arms
e. *poe piipii na piipii pak*
crying is the baby
f. *naa pe a tummuwali malo*
mama raises you in her arms
g. *pani nukku pa ki pani*
(refers to something that mama is doing while sitting)
h. *nokku we ye*
(refers to something that mama is doing while sitting)

i. *nana peka*
mama is sitting
j. *u kachi pa kina*
she is sitting in the hammock
k. *nana peka nai kucha ye*
mama is sitting while the little ones come
l. *pani pii nana pani poe tii kpaa um*
you (little girl) were crying
m. *puna tola piipii maloye*
my little girls
n. *machi tola o kanapi*
you will grow up and marry a grown-up boy
o. *na pe o tumnotali maloye*
you will grow up

Repetition of the word sequence *nana peka u kachi pa kine* is evident in (1c), (1i), and (1j). The very common words *poe pi, poe pii, poa tii,* and *poe piipii* (relating to crying) appear in lines (1b), (1d), (1e), and (1l). *Nana* (mama), a common initial word pattern, occurs in lines (1c), (1d), (1i), and (1k). Finally, *maloye* and *malo* are typically used to mark the end of a musical phrase or section in lines (1b), (1f), (1m), and (1o). Unfortunately, the translated text cannot capture the rhythm of the original because each Cuna syllable corresponds to a sung note.

Sakata (1987), in her work on the Hazara tribe in central Afghanistan, noted that Hazara women, who are excluded from musical activities in their male-dominated culture, distort or invent words in their lullabies, with the apparent goal of producing mellifluous sounds. This orientation toward "sound effects" likely accounts for the cross-cultural prevalence in lullabies of stereotyped syllables such as *loo-loo, lulla, nina, bo-bo,* and *do-do* (Brakeley, 1950; Brown, 1980). Liberal use of elongated vowels is also reported in lullabies sung by Mohave (Devereux, 1948), Arapaho (Hilger, 1952), Chippewa (Hilger, 1951), and Hopi (Sands & Sekaquaptewa, 1978) caregivers.

Moreover, lullabies seem to embody a number of properties associated with infant-directed speech in general and soothing infant-directed speech in particular. As noted earlier, soothing speech to infants includes low, falling contours, a narrow pitch range, and a gentle tone of voice (Fernald & Simon, 1984; M. Papoušek & Papoušek, 1981; M. Papoušek et al., 1991). In fact, the features of narrow pitch range, smooth repeating contours, and repetitive rhythms have been documented for Vietnamese (Cong-Huyen-Ton-Nu, 1979), North American Indian (Sands & Sekaquaptewa, 1978), Hazara (Sakata, 1987), and Cuna (McCosker, 1974) lullabies. Nevertheless, falling contours have not been mentioned in these descriptions. Like the early research on infant-directed speech, however,

most work on infant-directed song has been focused on textual content (i.e., the verbal message) rather than melodic form.

B. LULLABY IDENTIFICATION ACROSS CULTURES

Despite the functional distinctiveness of lullabies (i.e., their use in accelerating sleep), the question arises as to whether they are perceptually distinct to listeners unfamiliar with the culture. In the case of infant-directed speech, distinctiveness is considered a foregone conclusion, making experimental verification seem unnecessary. Trehub, Unyk, and Trainor (in press-a) evaluated this question with lullabies from different cultures. They paired 30 foreign lullabies with comparison songs (mostly adult songs) and asked adult listeners to identify the infant-directed song in each pair. The songs were selected from ethnomusicological materials recorded in diverse cultures and geographic regions, with none sung in a language familiar to the adult participants. Because the characteristically slow tempo of lullabies could provide an obvious cue, each comparison song was matched in tempo as well as language and cultural origin to the relevant lullaby. Adult listeners identified the lullabies significantly better than chance and their performance was independent of musical training (i.e., years of music lessons) and familiarity with the musical system (i.e., Western versus other).

Perhaps mellifluous or stereotyped syllables (Brown, 1980; Sakata, 1987), onomatopoeia (Curtis, 1921), and word reduplication (McCosker, 1974) accounted for the listeners' performance. To reduce or eliminate such influences, Trehub et al. (in press-a) content-filtered the songs by removing all frequencies above 500 Hz. The resulting recordings sounded muffled, with the words completely obscured but the melody line and many aspects of voice quality largely intact. Adult listeners were still able to identify the lullabies, indicating that information conveyed by melody or voice quality must provide distinctive cues.

To eliminate voice quality and other performance cues, Trehub et al. created a synthesized (piano-timbre) version of the melody line of the lullabies and comparison songs. This manipulation also removed all vocal ornamental devices such as trills and glides, which eliminated much of the "foreign" sound of the materials. Performance on these reduced versions of the songs was much less accurate, indicating the contribution of vocal quality or tone of voice to lullaby identification. Nevertheless, performance was still significantly correlated with performance on the original materials. Taken together, these findings are indicative of the contribution of melodic as well as word cues to the identification of lullabies. In short, soothing infant-directed song is perceptually distinct, at least for adult listeners.

The task was by no means easy, performance being at approximately 66%

correct on the original materials. The noteworthy finding, however, was the consistency of performance. For example, four of the lullabies (Creek Indian, Czech, Irish, Pygmy) were correctly identified by more than 85% of the listeners and four other lullabies (Chadian, Ecuadorian, Samoan, Ukranian) were consistently but incorrectly rejected. This distinct pattern of performance prevailed even in the filtered and synthesized versions. Perhaps adults had some conception of a lullaby, a pancultural stereotype or prototype, that they used to judge the songs they heard.

Other work with these lullabies and comparison songs (Unyk, Trehub, Trainor, & Schellenberg, 1992) has revealed that adult listeners judge lullabies to be significantly simpler than comparison songs whether these are presented in original, filtered, or synthesized versions. Moreover, the lullabies identified with greatest accuracy (in Trehub et al., in press-a) are rated as simpler than those identified with least accuracy.

Although simplicity in general distinguishes infant-directed speech and song from their adult-directed counterparts, little is known about the component features that contribute to the perceived simplicity of songs and their appropriateness for infants. For example, simple songs such as lullabies might have fewer changes in pitch direction (i.e., contour changes) per unit time. Moreover, soothing songs, like soothing maternal speech, might have a preponderance of descending contours.

In a musicological analysis of transcribed (i.e., written) versions of the lullabies and comparison songs, Unyk et al. (1992) evaluated a number of possible differentiating features such as median pitch, pitch range, phrase length, contour complexity, and descending contours. Surprisingly, none of these features, either singly or in combination, distinguished the lullabies from the other songs. Nevertheless, some features were reliably associated with adults' judgments of lullabies. In particular, songs with a greater proportion of descending contours were more likely to be judged as lullabies, as were those with fewer contour changes. These features may be components of adults' prototype or mental representation of lullabies (Trehub & Unyk, 1992).

What factors may have prevented listeners from achieving greater accuracy in lullaby identification? Perhaps the choice of materials obscured important structural as well as stylistic (i.e., performance) distinctions. The tempo of lullabies and comparison songs, no doubt a critical lullaby cue, had been deliberately equated, resulting in a likely underestimation of the cross-cultural recognizability of lullabies. In addition, Trehub et al. used recorded materials from other investigators (ethnomusicologists) and therefore had no access to information about the context of the performances. As a result, the context may have been functionally appropriate in some cases (e.g., a mother lulling her infant to sleep) and inappropriate in others (e.g., an individual simply responding to a researcher's request, perhaps in the absence of an infant). The essence of a lullaby may derive,

at least in part, from fine-tuning the performing style to the function and context. The English verb *lull* attests to the importance of function by referring to a particular manner of soothing by sounds or caresses. In this regard, Hilger (1952) described the Arapaho tribe's occasional use of traditional dance songs as "lullabies." Sakata (1987) used the term *functional lullabies* to distinguish songs that women sing to their infants from the *stylized lullabies* that men sing (often with instrumental accompaniment) about their infants. Sands and Sekaquaptewa (1978), noting the contribution of infant characteristics to the singer's performance, distinguished between soothing lullabies for cooperative infants and admonishing lullabies for reluctant sleepers.

Just as an infant's presence may be necessary for the full set of appropriate speech adjustments (Fernald & Simon, 1984; M. Papoušek, Papoušek, & Haekel, 1987), so the infant's presence and appropriate state may be critical for the appropriate realization of a lullaby. Whether this realization involves an appropriate tone of voice (e.g., soft, gentle) or an appropriate dynamic quality (e.g., lulling) remains to be determined. In any case, such questions must await comparisons between contextually appropriate and inappropriate lullabies.

C. MATERNAL SINGING: THE INFANT'S CONTRIBUTION

Some research on maternal singing offers further perspectives on infant-directed song in general and on contextual factors in particular. Trehub, Unyk, and Trainor (in press-b) recorded mothers of infants (12 months or less) as they informally sang a song of their choice. For half of the mothers, the infant was present, in which case they sang a song directly to the infant. For the other half, the mothers were instructed to sing as they normally would while alone. To minimize the singer's self-consciousness, the experimenter left the room before the recording session began and the mother started and stopped the recording equipment. After completing her song, the mother recalled the experimenter, who requested that she sing the same song again but in the opposite condition (e.g., without the infant if the infant was originally present). Once again, the mother controlled the equipment. A composite recording with pairs of excerpts from the two conditions (from each of 15 mothers) was presented to adult listeners, who were required to identify the infant-directed excerpt in each pair. Listeners were highly accurate on this task, achieving scores of about 91% correct. Instrumental analyses revealed that the singing to infants was characterized by higher pitch and slower tempo. Moreover, musically trained listeners rated the infant-directed singing as having greater rhythmicity, a softer tone of voice, more elongated vowels, and a "smiling" quality. Smiling is known to alter vocal tract characteristics and the resulting performance of singers (Fonagy, 1981; Sundberg, 1973) as it does for talkers (Laver, 1980).

Another group of mothers provided samples of simulated and actual infant-

directed singing, which led to about 77% correct identification of actual singing to infants. A few of the mothers were better "actors" than others, generating lively and convincing simulations. In any case, the infant's presence seems to alter features of the singer's performance in a way that enhances its identifiability.

Does culture-specific knowledge facilitate the identification of performance features associated with contextual appropriateness (i.e., an infant's presence)? To gain insight into this question, Trehub et al. (in press-b) recorded Indian mothers (in India) and North American mothers of Indian descent as they sang Hindi songs with their infant present or absent. North American listeners were less accurate in judging the context of these foreign songs (57% correct) than they were with North American singers (about 91% correct), but their performance significantly exceeded chance levels. Listeners of Indian origin performed better (71% correct) than native-born North Americans on these Hindi materials, but their performance was poorer than that of native-born North American listeners on English materials.

The choice of songs may have contributed to listeners' poorer performance on Hindi than on English songs. The English singers tended to choose play songs such as *Twinkle, Twinkle Little Star* or *Frère Jacques*. Such play songs were selected even when mothers began the session by singing on their own after being instructed to sing something that they usually sang. In contrast, the Hindi singers tended to select soothing or religious songs in both contexts so that the overall character of sung materials was very different. Perhaps soothing songs, with their smooth contours and slow tempo, offer less scope for performance variations than do playful songs. This difference may also characterize soothing versus playful speech, but no research has been done on this issue.

The cross-cultural variations in infant-directed speech may have their counterpart in infant-directed song, with arousing songs predominating in typical North American and European contexts and soothing songs predominating in other cultural contexts. In India, for example, mothers tend to remain with their infants until they fall asleep, often singing and otherwise comforting them. Many of the Indian mothers in the Trehub et al. (in press-b) sample reported similar practices and some even selected such a context for their recording. In contrast, North American infants typically sleep in a separate room and their caregivers often withdraw before they fall asleep. In fact, native-born North American mothers in the Trehub et al. (in press-b) study reported that they sang lullabies infrequently, if at all. Caudill and Weinstein (1969) observed more lullaby singing by Japanese than by North American mothers, although more recent reports (Sengoku, cited in Morikawa et al., 1988) reveal some "modernization" of Japanese caregiving, including a reduction in holding and rocking and a corresponding increase in talking.

D. PATERNAL SINGING

Do fathers show less distinctiveness in their songs to infants, as they do in their infant-directed speech (Fernald et al., 1989)? Work in progress by Trehub, Unyk, and their associates supports an affirmative answer. Fathers were recorded while singing to their infants or while simulating such singing. Unlike mothers, few sang standard children's songs. Instead, they sang popular or folk tunes, freely improvising the words or melodies, and often embedding the infant's name in their songs. Nevertheless, listeners had more difficulty differentiating actual from simulated paternal songs than they did with maternal songs. Were the fathers such excellent simulators that they misled the listeners or were they simply less "attuned" to their infants? On the basis of independent ratings of mothers' and fathers' singing style, less attunement seems to have been the case.

IV. Infants' Processing of Speech Sequences

Infant-directed speech has a potent effect on its intended audience. Such speech induces heightened affective responsiveness (Werker & McLeod, 1989), smiling (Wolff, 1963), vocalization (Mayer & Tronick, 1985; Stevenson, Ver-Hoeve, Roach, & Levitt, 1986), and vocal imitation (Lieberman, Ryalls, & Rabson, cited in Lieberman, 1984). Moreover, infants are rated as more likable when listening to infant-directed than to adult-directed speech (Werker & McLeod, 1989). This speech register also facilitates infants' differentiation of their mother's voice from that of a stranger (Mehler, Bertoncini, Barrière, & Jassik-Gerschenfeld, 1978) and their differentiation of phonetic contrasts in sequences of syllables (Karzon, 1985).

A. ATTENTIONAL PREFERENCES

Some researchers have systematically evaluated the relative attention-eliciting properties of natural or systematically altered infant- and adult-directed speech. In one variant of this procedure (Fernald, 1985), infants are presented infant-directed speech when they look at one loudspeaker and adult-directed speech when they look at another. Longer elective looking for one type of speech over the other is considered to reflect its greater attention-eliciting properties or infants' "preference" for that speech variety. Werker and McLeod (1989) added a visual component, creating audio–video displays of infant- and adult-directed speech. In another variant of the preference procedure (Cooper & Aslin, 1990), the presentation of sound is made contingent on looking at a single loudspeaker, with the two types of speech presented on alternating trials. In yet another

procedure (Pegg, Werker, & McLeod, 1992), looking time to both types of speech is compared in the context of an infant-controlled habituation design.

Despite procedural differences, the results have been consistent in revealing attentional preferences favoring infant-directed speech in newborns and 1-month-olds (Cooper & Aslin, 1990), 1.5-month-olds (Pegg et al., 1992), 4- to 5-month-olds (Fernald, 1985; Werker & McLeod, 1989), and 7- to 9-month-olds (Werker & McLeod, 1989). Moreover, infant-directed speech produced by fathers generates an attentional preference but does not noticeably heighten affect (Pegg et al., 1992; Werker & McLeod, 1989).

Some progress has been made in isolating the features responsible for the perceptual salience of infant-directed speech. For example, 4-month-olds' preference for infant-directed speech was found to persist with speech samples filtered to remove the lexical content but not the prosody (Fernald, 1985). A similar attentional bias prevailed for synthesized contours that kept the pitch and temporal patterning intact but not for those that maintained temporal or amplitude information divorced from pitch patterning (Fernald & Kuhl, 1987). The situation seems to be somewhat different for 1-month-olds. Researchers who have used content-filtered infant- and adult-directed speech or synthesized pitch contours have not found the differential attentiveness obtained with natural speech samples (Cooper, in press). Possibly, a number of features such as pitch, pitch contour, and temporal and amplitude patterning act jointly to increase the impact of infant-directed speech. The less mature the listener, the more features may be required to trigger differential responsiveness. With older infants, however, some features may be effective singly (e.g., high pitch) and others only in combination. Nevertheless, the presence of multiple features no doubt increases the magnitude of the response.

B. AFFECTIVE PREFERENCES

In an effort to discern possible "meanings" in infant-directed speech for infant listeners, other researchers have examined differential affective responses to "positive" and "negative" infant-directed speech. In one such study (M. Papoušek, Bornstein, Nuzzo, Papoušek, & Symmes, 1990), 4-month-old infants were presented simulated pitch contours characteristic of approving and disapproving (prohibitive) infant-directed speech. Each simulated utterance was produced by extending a neutral vowel over a typical infant-directed contour. In line with the investigators' predictions, infants were more visually attentive in the context of the approving contours.

In a related investigation (Fernald, 1992), 4-month-olds listened to infant-directed utterances of approval or prohibition in familiar (English) or unfamiliar (German, Italian, Greek, Japanese) languages. Infants exhibited more positive affect while listening to approving than to prohibiting utterances except for

Japanese utterances. The Japanese approvals and prohibitions were found to embody a narrower pitch range and were rated by adults as less intense than their counterparts in the other languages.

Such findings as well as the attentional biases for natural, filtered, or synthesized infant-directed speech have led to an emphasis on the fundamental frequency (i.e., pitch) characteristics of infant-directed speech. Discussions of these results have tended to center on heightened pitch and expanded pitch range despite the fact that the pitch contours in question were always presented with temporal patterning appropriate to the context (infant- or adult-directed). Without separate manipulations of temporal and pitch patterning, researchers cannot disentangle their respective contributions.

In the two studies of approving and disapproving infant-directed speech (Fernald, 1992; M. Papoušek et al., 1990), the temporal characteristics of these utterances contrasted markedly. The approving utterances embodied gradual modulations of pitch and loudness as opposed to the abrupt changes of disapproving utterances. Nevertheless, these researchers and others have tended to focus on pitch patterning at the expense of temporal patterning when they refer to the melodic contours or melodies of infant-directed speech.

C. TEMPORAL PATTERNING AND ATTENTIONAL PREFERENCES

The temporal patterning of infant-directed speech has received more attention from researchers interested in the relations between prosodic features and syntactic structure (e.g., Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright-Cassidy, Druss, & Kennedy, 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989). For example, clause boundaries tend to be marked by prosodic features such as final syllable lengthening and pitch change. Might prosodic marking be enhanced in infant-directed speech and would prelinguistic infants be sensitive to such prosodic features? To address these questions, Hirsh-Pasek et al. (1987) and Kemler Nelson et al. (1989) altered recordings of speech by inserting pauses at clause boundaries (i.e. between clauses) or at other locations (i.e., within clauses). The former manipulation maintained the structural integrity of clauses and the latter altered the internal temporal structure. Infants (7- to 10-month-olds) exhibited a preference for intact clauses when the speech was infant directed but no preference when it was adult directed.

A number of implications follow from these findings. First, prelinguistic infants process temporal pattern cues in complex speech sequences, particularly when such cues are linked to other important prosodic features. Second, infants are sensitive to prosodic cues associated with syntactic (i.e., clause) structure in infant-directed speech. Third, cues to at least some aspects of syntactic structure are enhanced in infant-directed speech. Such links between salient prosodic cues and critical syntactic features could facilitate infants' parsing of the speech

stream and, by extension, their acquisition of language (Kemler Nelson et al., 1989; Morgan, Meier, & Newport, 1987).

V. Infants' Processing of Musical Sequences

Research on infants' processing of speech sequences has provided insights into the impact of infant-directed speech on its intended audience. Unfortunately, no researchers in the musical domain have examined the effects of infant-directed song on infant listeners; however, some information is available on infants' perceptual capabilities with respect to music in general. Such information permits preliminary conjectures about infants' potential for responding differentially to infant- and adult-directed music.

A. ATTENTIONAL PREFERENCES

One investigation of infants' temporal processing of musical sequences was a direct analog of the Hirsh-Pasek et al. (1987) and Kemler Nelson et al. (1989) studies of prosodic cues to clausal structure. Krumhansl and Jusczyk (1990) examined infants' sensitivity to phrase structure in music. To do so, they created two temporally altered versions of Mozart minuets. In one, they inserted pauses between musical phrases, thereby maintaining the internal temporal structure of such phrases. In the other, they altered the phrase structure by inserting pauses within phrases. Infants (4 and 6 months of age) showed greater attentiveness for the minuets with intact phrases than for those with altered phrases, indicating their sensitivity to musical phrase structure. An analysis of the phrase endings of these minuets revealed features similar to those that mark clause boundaries in speech. For example, the last melody note of the phrases tended to be lengthened, much like the last syllable in clauses. Also, musical phrases tended to end with a drop in pitch. Changes in pitch, whether increases or decreases, tend to characterize clause endings. Thus, infants are sensitive to the prosodic integrity of phrases in music and clauses in speech.

In contrast to Krumhansl and Jusczyk's (1990) research on attentional biases or preference, most other research on musical processing in infancy has been focused on the perception of various aspects of music, including pitch and temporal patterning (for reviews see Trehub, 1990; Trehub & Trainor, 1990, 1993). For the present purposes, this work is considered in relation to the following broad questions. What features of music are salient for infant listeners? Are some patterns processed more readily than others? Is "premusical" infants' processing of music culture independent?

B. METHODOLOGICAL STRATEGY

The general procedural approach in these studies is to present melodies (up to 10 notes) that cannot be remembered in their entirety. In this way, the details actually retained are informative about infants' characteristic processing strategies. For example, infants' retention of the exact pitches of the first or last few notes would be indicative of a local processing strategy; however, their retention of general pattern information (e.g., relations between the pitches or durations of notes) as opposed to specific details (e.g., exact pitches or durations) would be indicative of a global processing strategy. This procedural approach is implemented with an operant discrimination design (see Trehub, Bull, & Thorpe, 1984; Trehub, Thorpe, & Morrongiello, 1987). Specifically, infants 6 months and older are presented a standard pattern that is repeated continuously from a loudspeaker to one side and are trained to respond (i.e., turn) when a comparison pattern is substituted for the standard. Correct responses (i.e., turns to the comparison pattern) lead to reinforcement in the form of brief presentations of an animated toy; incorrect responses (false positives or incorrect rejections) have no consequences. By judicious selection of standard and comparison melodies, insight can be gained into infants' processing strategies in general and the specific features that evoke discriminative responding.

Typically, standard and comparison melodies are presented in different pitch registers (i.e., transposed) to study infants' processing of pattern information rather than absolute pitch. In fact, infants treat exact transpositions as equivalent, just as adults do (Chang & Trehub, 1977; Trehub et al., 1984). They do not simply fail to discriminate the relevant differences but rather focus on relational as opposed to absolute pitch information (Trehub, 1990; Trehub & Trainor, 1990).

C. PITCH PATTERNING

The results of various studies of melody perception have revealed infant performance that is surprisingly adult-like (Trehub & Trainor, 1993), with some features of melodies being much more salient than others. For example, a comparison melody that embodies a change in contour (i.e., the pattern of directional changes in pitch) relative to the standard melody is almost invariably responded to differentially by infant listeners (Trehub et al., 1984; Trehub, Thorpe, & Morrongiello, 1985); however, a comparison melody that embodies new notes but has the same pitch contour will likely be responded to in the same way as the original (Trehub et al., 1984, 1987). For infants, then, pitch-contour processing predominates in the perception of musical as well as speech patterns. Adults also treat contour as a highly salient feature of unfamiliar melodies (Dowling, 1978).

Another feature that influences infants' response to melodies is pitch range. When the pitch range of a comparison melody differs from that of the standard, infants are likely to respond even if relational pitch information is preserved (Trehub et al., 1984, 1985).

D. TEMPORAL PATTERNING

In the aforementioned studies, the temporal configuration (i.e., rhythmic patterning) was held constant so that the pitch configuration could be assessed independently. Other research has established that infants also respond to melodies or tone sequences on the basis of their rhythmic or temporal structure (i.e., relative durations of notes) rather than the absolute durations of component notes (Trehub & Thorpe, 1989). Just as the identity of a melody is independent of specific pitches, so is it independent of any specific tempo, for infant listeners as well as adults.

Infants also group or parse the component notes of patterns in much the same way as adults. For example, with patterns consisting of three tones of one type followed by three tones of another type (schematic structure: XXXOOO), infants detected pauses inserted within groups of similar notes (e.g., XXXO OO: structure-disrupting changes) more readily than pause inserted between groups of similar notes (e.g., XXX OOO: structure-conserving changes) (Thorpe & Trehub, 1989; Thorpe, Trehub, Morrongiello, & Bull, 1988). The propensity to group sounds in this way is likely implicated in infants' "preference" for intact phrases in music (Krumhansl & Jusczyk, 1990) and for intact clauses in speech (Kemler Nelson et al., 1989).

In short, infants readily encode information about pitch contour, pitch range, and temporal patterning from musical sequences. Thus, the very same features that predominate in infant-directed speech and influence infant attention toward such speech emerge as salient features of musical sequences. In all likelihood, infants use the same or very similar perceptual organizational devices in their processing of complex auditory patterns, whether speech or music (Trehub, 1989; Trehub & Trainor, 1993). Pitch contour and temporal patterning, for example, provide reasonable means for parsing the speech or musical stream into chunks appropriate for further processing.

E. MELODY AND CONTOUR

A number of researchers (e.g., Fernald, 1989, 1992; M. Papoušek & Papoušek, 1981; M. Papoušek et al., 1990, 1991) have referred to the typical pitch contours of infant-directed speech as melodies. This characterization ignores the usual distinctions between contour and melody. In musical parlance, contour is a general feature that captures directional pitch movement without regard to the

extent of such movement (Dowling, 1978; Dowling & Harwood, 1986). Descriptive terms like *rising, falling,* and *rise–fall* or bell-shaped contours are consistent with this usage. Melody, in contrast, is usually defined in terms of the exact pitch intervals and rhythmic relations between notes in the context of a musical system. Even a simple rising contour can differ by consisting of *doh re mi* or *doh mi soh,* for example. Designating the latter pattern as expanded ignores the fact that it is a different melody or tune, despite the ascending sequence of three notes in both cases. Similarly, relative durational differences between notes generate different melodies.

F. "GOOD" AND "BAD" MELODIES

As noted earlier (Section V,C), a number of investigations revealed that infants are primarily contour processors. This finding does not imply, however, that they are exclusively contour processors. As will become clear, for example, not all bell-shaped melodies are equivalent for infant listeners. Some are "good" in the sense that they are readily processed and others "bad," being processed with greater difficulty (Trehub & Trainor, 1993). Progress in delineating the critical features of such good and bad melodies might suggest new avenues for research on infant-directed speech.

Every language has an inventory of component sounds, approximate boundaries of sound categories, and rules for combining sounds. Correspondingly, every musical system has an inventory of musical pitches (notes) organized into scales as well as various conventions for arranging these notes into acceptable sequences. Presumably, constraints of the auditory system have had some influence on the selection and arrangement of sounds in music, as in speech (Stevens & Keyser, 1989). Some idiosyncratic or culturally based rules are probably also operative. If any natural musical rules or features can be identified, such features might be relevant in the early processing of speech prosody.

Extensive research has been conducted on the music processing skills of adult listeners, including those who have received incidental musical exposure rather than formal training (for reviews, see Deutsch, 1982; Dowling & Harwood, 1986; Handel, 1989; Krumhansl, 1990). This research, however valuable, cannot provide insights into natural rules or processing proclivities for a number of reasons. Principal among these is the long exposure to a musical system that would effectively mask the distinction between naturally good melodies and those that become good (i.e., easily processed) by means of such exposure. Infants, however, would have limited exposure to the music of their culture and thus are unlikely to have learned a great deal about its idiosyncrasies. In this sense, they are reasonable candidates for evaluating good patterns. Patterns that infants would process more readily than others could be considered good or as embodying good features. Learning is not necessarily excluded. The only re-

quirement is that processing priority for such patterns be exhibited in early life. If learning is implicated, this type of learning may be innately guided (Jusczyk & Bertoncini, 1988; Marler, 1990).

A number of investigators have confirmed that Western adults exhibit enhanced processing (e.g., more detailed encoding, better retention, preference) for melodies that conform to the musical conventions of their culture (Cuddy, Cohen, & Miller, 1979; Krumhansl & Keil, 1982; Watkins, 1985). This effect might stem from familiarity with such conventions or from naturally or inherently good features of such patterns (or both).

Researchers have identified some melodies that seem to be inherently good in the sense that Western infants process them in greater detail than is the case for other similar melodies. Of particular interest is the fact that these good melodies embody rules or conventions that are central to Western music. The investigations in question had standard and comparison melodies with bell-shaped (rise–fall) contours, the comparison melody differing minimally from the standard (i.e., a semitone in one position only). As usual, the standard and comparison melodies were presented in transposition, precluding the use of absolute pitch cues (see Fig. 2). When the standard melody was created from the principal notes of the Western major scale—notes considered important from a music-theoretic perspective (Piston, 1969) and from adult listeners' perspective (Krumhansl, 1983)—infants went beyond their usual contour processing strategy, and encoded and retained information about intervals (i.e., exact pitch relations between adjacent notes) (Cohen, Thorpe, & Trehub, 1987; Trehub, Thorpe, & Trainor, 1990; Trainor, 1991). When, however, the standard melody had one or more "wrong" notes in terms of Western scale structure (Cohen et al., 1987; Trainor, 1991; Trehub, Thorpe, & Trainor, 1990) or when it deviated from that structure in more substantial ways (Trehub, Thorpe, & Trainor, 1990), infants did not encode and retain interval information. These findings are consistent with the notion that aspects of Western major scale structure have their origins in universal constraints on auditory pattern processing. In other words, such aspects may be inherently good.

These findings do not imply that Western major scale structure is superior to alternative scale structures (e.g., those in other cultures). Rather, the notion is simply that Western major scale structure exemplifies one of many possible instances of good form. Scales that are fundamental to other cultures are also likely to exemplify good form. Accordingly, infants should exhibit comparable enhanced processing for melodies based on such foreign scales. In fact, research by Lynch, Eilers, Oller, and Urbano (1990) supports this contention. They evaluated the ability of Western 6-month-olds and adults (musically untrained) to detect very subtle pitch changes to simple rise–fall melodies consisting of notes from the Western major scale or from the Javanese *pélog* scale. Not surprisingly, adults performed better on the major than on the *pélog* melodies. Infants also

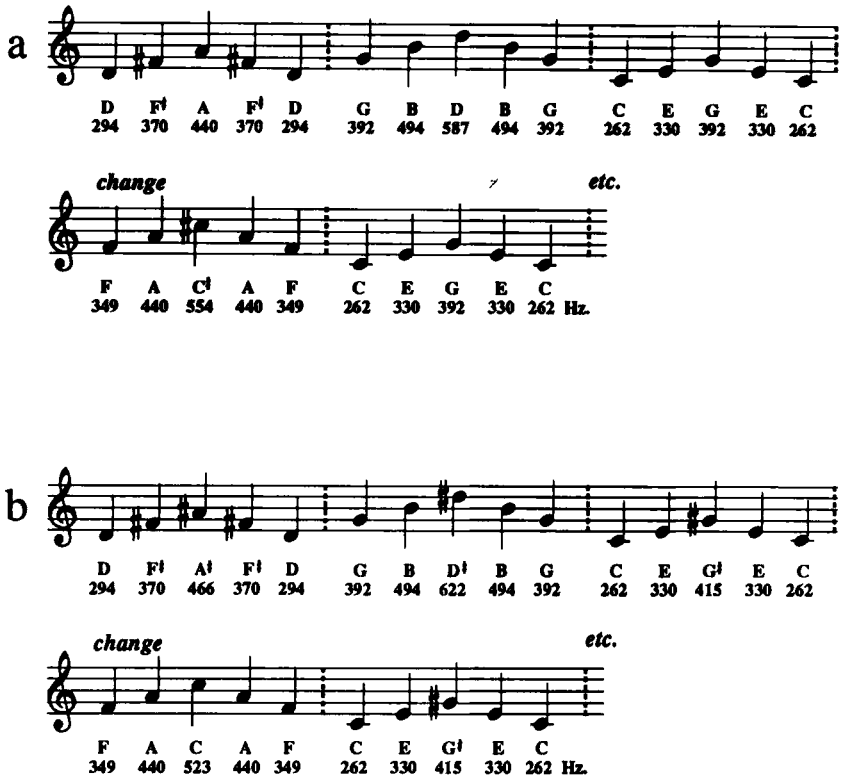

Fig. 2. Each example depicts three successive presentations of the standard melody followed by a change and subsequent return to the standard melody. The upper melody (a) conforms to Western major scale structure; the lower melody (b) has one "wrong" note (the third).

detected changes to the melodies but, unlike adults, their performance was equivalent for the two structurally distinct scales types. This latter finding suggests that both scales begin on an equal footing for musically unacculturated listeners but that culture-specific experience enhances one relative to the other. Similar arguments have been advanced for the speech sounds of different languages (Best, McRoberts, & Sithole, 1988; Burnham, Earnshaw, & Clark, 1991; Werker & Lalonde, 1988).

What features of the major scale might contribute to this relative ease of processing? Its most prominent notes (e.g., first, third, and fifth notes) in terms of frequency of occurrence, perceptual ratings, and music-theoretic considerations (see Krumhansl, 1990) are related by simple (small-integer) frequency ratios. Such ratios may form the basis of naturally salient or minimally dissonant (rough) intervals (see Burns & Ward, 1982). The perfect fifth interval, which

relates the first and fifth notes of the major scale, approximates the ratio of 3 : 2, which is considered the second most consonant (i.e., least dissonant) interval (the most consonant being the octave with a ratio of 2 : 1). Of interest is the fact that the 3 : 2 ratio is prominent in the overtones of vowels and many other natural sounds. Is it a mere coincidence that perfect fifth intervals figured prominently in the Western melodies for which infants exhibited interval processing (Cohen et al., 1987; Lynch et al., 1990; Trainor, 1991; Trehub, Thorpe, & Trainor, 1990)? Despite the much heralded differences in the component notes of different scales (Harwood, 1976), perfect fifth intervals are prominent cross-culturally, particularly in vocal music (Kolinski, 1967; Nettl, 1956).

Infants' enhanced performance (i.e., interval processing) for melodies based on the Western major (Cohen et al., 1987; Trainor, 1991; Trehub, Thorpe, & Trainor, 1990) and Javanese *pélog* (Lynch et al., 1990) scales is indicative of culture-independent processing. Further support for culture-free music processing in infancy has been provided by Trainor and Trehub (1992a). They presented infants and adults a 10-note standard melody based on the major scale and a comparison melody with only one note altered relative to the standard. The altered note was either consistent with Western musical rules or in violation of these rules (i.e., within or outside the prevailing key). For adults, a small pitch change that violated musical conventions was much easier to detect than a considerably larger pitch change that followed cultural conventions. For infants, however, performance was equivalent for both changes. Of particular interest was the fact that infants' level of performance on the rule-conserving change significantly exceeded that of adults. Essentially, adults' tacit knowledge of Western musical structure obscured a pitch difference that was perceptible for infant listeners. This finding parallels infants' ability to discriminate some foreign speech contrasts that are difficult for adults (Trehub, 1976; Werker & Tees, 1984; Werker & Lalonde, 1988).

G. GENERAL PERSPECTIVES

What can we conclude about musical processing in infancy? At the very least, infants have the prerequisite skills for processing musical input directed to them. They can parse the musical stream; they engage in relative pitch and temporal processing; and they characteristically extract pitch contours and rhythms from musical sequences. For certain patterns, so-called good melodies, they can encode and retain subtle details such as the precise relations between adjacent notes (i.e., intervals).

Infants' adult-like perception of musical patterns implies that many aspects of music processing are not arbitrary. Instead, such processing seems to engage basic principles of perceptual organization that are operative early in life. The present conception challenges the prevailing view that long-term exposure to

music promotes the extraction of pitch and temporal regularities and the internalization of musical conventions (Krumhansl, 1990; Jones, 1990). No doubt some aspects of musical structure are arbitrary or conventional (Trainor & Trehub, 1992a), but these may be relatively few. The views espoused here clash with those of contemporary composers and music theorists (e.g., Boulez, 1971; Schoenberg, 1975), who contend that audiences, with comparable exposure, would be as comfortable with 12-tone compositions as with music from Beethoven or The Beatles.

VI. Music and Speech: Present Perspectives and Future Prospects

A. CAREGIVERS' SPEECH AND SONG

A number of intriguing parallels have been found between speech and song to infants. Both seem to involve significant adjustments in relation to standard speech and song, adjustments that go beyond the speaker's or singer's awareness. Among such adjustments are simple contours, elongated vowels, and emotional expressiveness. Speech and song alike have variants linked to the infant's state and perhaps to the caregiver's state, as well. Thus, speech to capture and hold infants' attention has its counterpart in play songs, and soothing speech has its counterpart in lullabies. Further, when speech and song are perceived to be of the soothing variety, descending contours prevail. To the extent that speech and song have meaning for infant listeners, these meanings are likely affective and nonarbitrary.

The parallels between speech and song to infants are especially impressive in view of the fact that adult-directed speech has no musical analog (except for artificial performing contexts such as opera). Although singing is not conventionally directed to individuals, caregivers from time immemorial have considered one-on-one singing a necessary or desirable part of their interactive repertoire with infants.

Lullabies and play songs can be distinguished from other sung materials within a culture but they are similar in many respects to these other materials. All songs from a culture, regardless of their audience, conform to the same musical system and, consequently, share a number of features that distinguish them from the songs of other cultures. In descriptions of speech, the emphasis has been on differences between infant- and adult-directed varieties as well as cross-cultural similarities in speech to infants (e.g., Fernald et al., 1989; Grieser & Kuhl, 1988). The cross-cultural similarities and within-culture differences may have been overstated, however. If pitch contours constitute the essence of speech to infants and such contours are universal, then infant-directed speech samples from

one culture should be more similar to those of other cultures than to adult-directed samples from the same culture (all samples content filtered, of course). In other words, language identity should be obscured in the prosody of speech to infants. This outcome is unlikely to be the case. Rather, cues to language identity probably coexist with cues to infant-directedness. In any event, the relative contributions of these factors should be empirically established so that the phenomenon of infant-directed speech can be understood more fully.

Differences between infant- and adult-directed speech may also have been exaggerated. Typically, emotionally charged infant-directed speech is compared with emotionally neutral adult-directed speech. For example, Fernald and Simon's (1984) speech samples were derived from mothers talking to their 4-month-old infants or to the adult interviewer. Surely, features attributable to heightened (positive) affect must be separated from those attributable to infant-directedness. Samples of speech from new lovers might shed light on this issue. For example, happiness and joy in adult-directed speech are known to be associated with high pitch, large pitch variability, and an upward pitch contour, and pleasantness is known to be associated with low pitch and a downward pitch contour (Scherer & Oshinsky, 1977).

Another source of artifactual differences may be the stereotyped repertoire of infant-directed utterances. Stereotyped messages pose few cognitive demands on the speaker compared with the requirements of planning and executing syntactically, semantically, and phonologically appropriate messages for adults. As noted earlier (Section II,C), the prosody of stereotyped utterances to adults (e.g., *Hi! How are you? Great!*) may resemble infant-directed prosody more than typical adult-directed prosody. Perhaps highly stereotyped or rehearsed adult-directed messages would reveal the contribution of processing demands to prosodic form. Everyday experience clearly shows that actors, public speakers, and lecturers exhibit more pitch and loudness modulation than those with less experience or skill. Does this difference arise from the speaker's relative ease with the content, from effective simulation of an appropriate emotional state, or from sensitivity to the listener's preferences and needs?

Perhaps some vocal adjustments are triggered by the listener's level of verbal comprehension. For example, Hirsh-Pasek and Treiman (1982) observed speech of the infant-directed variety in adults' interaction with pets (termed *doggerel* by the authors). Do these adjustments arise from the listener's verbal intelligence (i.e., low), the speaker's feelings about the listener (e.g., warm), or both?

Disentangling the contributions of these various factors should lead to clearer and more informed distinctions between speech addressed to infants and adults. Such distinctions might shed new light on the origin and significance of common features in infant-directed speech and song.

One potentially important aspect of infant-directed speech, notably its voice quality, has not received the attention it merits. Anecdotal accounts include

references to "gentle," "soothing," or even "loving" tones of voice, but no systematic descriptions of acoustic or perceptual features relating to voice quality are available. In view of the known links of voice quality to emotional state (Frick, 1985; Scherer, 1981), greater specification of the spectral structure of infant-directed speech and song would be useful. This enterprise could be guided by research on the spectral structure of adult-directed speech, its relation to the speaker's emotional state, and the listener's decoding of that state (Laver, 1980; Lieberman & Michaels, 1962; Newman & Emanuel, 1991; Scherer & Oshinsky, 1977).

B. INFANTS' SPEECH AND SONG PROCESSING

The organization of speech into clauses and of music into phrases does not depend on the listener's comprehension of the speech content or familiarity with culture-specific conventions (Kemler Nelson et al., 1989; Krumhansl & Jusczyk, 1990). Instead, adult-like grouping processes seem to be available in infancy to guide the segmentation or parsing of auditory sequences (Demany, 1982; Thorpe & Trehub, 1989; Thorpe et al., 1988). Common mechanisms seem to operate for speech and musical sequences, highlighting important constituents in both cases (Krumhansl, in press; Trehub & Trainor, 1993). The success to date in identifying infants' grouping and segmentation strategies with respect to musical sequences constitutes an effective argument for further research of this nature.

The identification of linguistically relevant prosodic cues in speech to infants should not obscure the fact that these cues are also present in speech to adults. Intonation (patterns of pitch, loudness, and timing) in speech to adults has a multiplicity of functions including the portrayal of emotions, the identification of grammatical constituents, and the selective emphasis of particular meanings (Handel, 1989, p. 448). Good or expressive speakers, those who sound natural, tend to produce well-modulated speech with clear segmentation cues. Such cues may allow the listener to decode spoken messages with less effort than is necessary for messages with less prominent segmentation cues. A reasonable strategy would be to evaluate the processing advantages, even for adults, that arise from prominent prosodic features in infant-directed speech. An important need, moreover, is to establish whether differences in the incidence of segmentation cues in infant- and adult-directed speech are real or simply a by-product of inappropriate sampling of adult-directed speech, as outlined earlier (Section VI,A).

When prosodic features in infant-directed speech are viewed as facilitators of language acquisition (Kemler Nelson et al., 1989), the implication is that such features are unique either to infant-directed speech or to speech in general. Not only are these features present in adult-directed speech and music, but they also mark important boundaries in both (Handel, 1989; Krumhansl & Jusczyk, 1990). A reasonable conclusion, then, is that such features are domain general rather

than domain specific. Further, prosodic cues to linguistic and musical boundaries are perhaps more transparent in simple speech and musical sequences (e.g., those directed to infants) than they are in more complex sequences (e.g., those directed to adults). Few would dispute the notion that simple stimuli can be processed more readily than more complex stimuli; however, few would generate identical sets of defining criteria for simplicity. In this light, the ongoing debate about the utility of infant-directed speech for later language acquisition (e.g., Gleitman, Newport, & Gleitman, 1984; Kemler Nelson et al., 1989; Murray et al., 1990) may not be fruitful. Instead, a more profitable approach may be to establish the motivational basis for infant-directed speech adjustments and to delineate in greater detail the immediate effects on infants.

Infants exhibit an attentional "preference" for infant- over adult-directed speech (e.g., Cooper & Aslin, 1990; Fernald, 1985) and for approving over disapproving infant-directed speech (M. Papoušek et al., 1990) together with greater affective responsiveness to such signals (Fernald, 1992; Werker & McLeod, 1989). Comparable issues in infant-directed song have not been explored to date. One could, however, determine whether infants prefer lullabies over adult songs and whether such a preference (if present) is demonstrable only in sleepy or fussy infants. Would awake, alert infants prefer play songs over lullabies and also over adult songs? Do lullabies actually induce sleep more rapidly (e.g., briefer latency of sleep onset) than other songs?

If voice quality differs as a function of the speaker's or singer's emotional state, are such differences perceptible to infant listeners? Infants can differentiate and categorize the timbre or quality of single sounds (Clarkson, Clifton, & Perris, 1988; Trehub, Endman, & Thorpe, 1990) but whether they can discriminate timbre in sound sequences and whether voice timbre is salient for them are unknown. Those who argue that the production of emotional voice qualities is biologically based (see Frick, 1985) maintain that their perception has a similar basis, as is the case for emotional qualities associated with facial expression (Izard, 1977) and patterns of touching (Clynes, 1977). In any event, the study of infants' discrimination of emotionally relevant voice qualities and their preference for particular qualities is clearly warranted.

One finding arising from the musical domain, that of enhanced processing for so-called good melodies (e.g., Trehub, Thorpe, & Trainor, 1990), has no obvious parallel in infant-directed speech. To date, descriptions of maternal "melodies" in speech have not proceeded beyond the specification of their contours (e.g., Fernald et al., 1989) or distinctions between smooth and abrupt pitch and loudness transitions (e.g., Fernald et al., 1989; M. Papoušek et al., 1990). Perhaps the pitch ratio of 3 : 2, which seems to characterize good melodies, also figures prominently in typical infant-directed speech (Trehub & Trainor, 1993). This ratio could be highlighted in a number of ways. Its prominence in the simultaneous (i.e., harmonic) components of vowels could be enhanced by their

elongation, which also occurs in infant-directed speech and song, or by heightened pitch, which occurs in both. At a more speculative level, the relations between significant sounds in speech such as stressed words or syllables could embody the pitch ratio of 3 : 2. If so, Fernald's (1989) claim that "the melody is the message" would take on new meaning.

REFERENCES

Anderson, E. S. (1986). The acquisition of register variation by Anglo-American children. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (pp. 153–166). Cambridge: Cambridge University Press.

Aslin, R. N., Pisoni, D. B., & Jusczyk, P. W. (1983). Auditory development and speech perception in infancy. In P. H. Mussen (Ed.), *Handbook of child psychology*. Vol. II: *Infancy and developmental psychology* (pp. 573–687). New York: Wiley.

Bebey, F. (1969). *African music: A people's art*. Westport, CT: Lawrence Hill.

Beebe, B., Feldstein, S., Jaffe, J., Mays, K., & Alson, D. (1985). Interpersonal timing: The application of an adult dialogue model to mother–infant vocal and kinesic interactions. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 217–247). Norwood, NJ: Ablex.

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.

Bolinger, D. (1970). Relative height. In D. Bolinger (Ed.), *Intonation: Selected readings* (pp. 137–157). Hammondsworth: Penguin.

Boulez, P. (1971). *Boulez on music today* (S. Bradshaw & R. R. Bennett, trans.). London: Faber & Faber.

Brakeley, T. C. (1950). Lullaby. In *Standard dictionary of folklore, mythology, and legend* (pp. 653–654). New York: Funk & Wagnalls.

Brown, M. J. E. (1980). Lullaby. In *The new Grove dictionary of music and musicians* (pp. 313–314). London: Macmillan.

Burnham, D. K., Earnshaw, L. J., & Clark, J. E. (1991). Development of categorical identification of native and non-native bilabial stops: Infants, children and adults. *Journal of Child Language, 18*, 231–260.

Burns, E. M., & Ward, W. D. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (pp. 241–269). Orlando, FL: Academic Press.

Cass-Beggs, B., & Cass-Beggs, M. (1969). *Folk lullabies*. New York: Oak Publications.

Caudill, W. A., & Weinstein, H. (1969). Maternal care and infant behavior in Japan and America. *Psychiatry, 32*, 12–43.

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

Clarkson, M. G., Clifton, R. K., & Perris, E. E. (1988). Infant timbre perception: Discrimination of spectral envelopes. *Perception & Psychophysics, 43*, 15–20.

Clynes, M. (1977). *Sentics, the touch of emotions*. New York: Doubleday.

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.

Cong-Huyen-Ton-Nu, N.-T. (1979). The functions of folk songs in Vietnam. In J. Blacking & J. W. Kealiinohomoku (Eds.), *The performing arts: Music and dance* (pp. 141–151). The Hague: Mouton.

Cooper, R. P. (in press). The effect of prosody on young infants' speech perception. *Advances in Infancy Research.*

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584–1595.

Crystal, D. (1973). Non-segmental phonology in language acquisition: A review of the issues. *Lingua, 32,* 1–45.

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

Curtis, N. (1921). American Indian cradle songs. *The Musical Quarterly, 1,* 549–558.

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

Deutsch, D. (1982). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music* (pp. 99–134). New York: Academic Press.

Devereux, G. (1948). The Mohave neonate and its cradle. *Primitive Man, 21,* 1–18.

Dixon, S., Tronick, E., Keefer, C., & Brazelton, T. B. (1981). Mother–infant interaction among the Gusii of Kenya. In T. M. Field, A. M. Sostek, P. Vietze, & P. H. Liederman (Eds.), *Culture and early interactions* (pp. 149–168). Hillsdale, NJ: Erlbaum.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85,* 341–354.

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

Dunn, J., & Kendrick, C. (1982). The speech of two- and three-year-olds to infant siblings: "Baby talk" and the context of communication. *Journal of Child Language, 9,* 579–595.

Emanuel, F. W., & Smith, W. F. (1974). Pitch effects on vowel roughness and spectral noise. *Journal of Phonetics, 2,* 247–253.

Ferguson, C. (1964). Baby talk in six languages. *American Anthropologist, 66,* 103–114.

Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, C. Garvey, & R. Golinkoff (Eds.), *The origins and growth of communication* (pp. 5–29). Norwood, NJ: Ablex.

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. (1992). Meaningful melodies in mother's speech to infants. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental aspects* (pp. 262–282). Cambridge: 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., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27,* 209–221.

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 Development, 16,* 477–501.

Ferrier, L. J. (1985). Intonation in discourse: Talk between 12-month-olds and their mothers. In K. Nelson (Ed.), *Children's language* (Vol. 5, pp. 35–60). Hillsdale, NJ: Erlbaum.

Fonagy, I. (1981). Emotions, voice, and music. *Research Aspects on Singing, 33,* 51–79.

Francès, R. (1958). *La perception de la musique.* Paris: Vrin.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97,* 412–429.

Garnica, O. K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 66–88). Cambridge: Cambridge University Press.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language, 11,* 43–79.

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.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

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

Hilger, M. I. (1951). Chippewa child life and its cultural background. *Bureau of American Ethnology Bulletin, 146.*

Hilger, M. I. (1952). Arapaho child life and its cultural background. *Bureau of American Ethnology Bulletin, 148.*

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright-Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26,* 269–286.

Hirsh-Pasek, K., & Treiman, R. (1982). Doggerel: Motherese in a new context. *Journal of Child Language, 10,* 23–37.

Izard, C. E. (1977). *Human emotions.* New York: Plenum.

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.

Jocić, M. (1978). Adaptation in adult speech during communication with children. In N. Waterson & C. Snow (Eds.), *The development of communication* (pp. 159–171). Chichester, England: Wiley.

Jones, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology, 9,* 193–228.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. E. Ferguson, L. Menn, & C. Stoll-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17–64). Timonium, MD: York Press.

Jusczyk, P. W., & Bertoncini, J. (1988). Viewing the development of speech perception as an innately guided learning process. *Language and Speech, 31,* 217–238.

Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology, 39,* 326–342.

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,* 55–68.

Knower, F. H. (1941). Analysis of some experimental variations of simulated vocal expressions of the emotions. *Journal of Social Psychology, 14,* 369–372.

Kolinski, M. (1967). Recent trends in ethnomusicology. *Ethnomusicology, 11,* 1–24.

Krauss, R. M., Curran, N. M., & Ferleger, N. (1983). Expressive conventions and the cross-cultural expression of emotion. *Basic and Applied Social Psychology, 4,* 295–305.

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

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

Krumhansl, C. L. (in press). Grouping processes in infants' music perception. In J. Sundberg, L. Nord, & R. Carl (Eds.), *Grouping in music.* Stockholm: Royal Swedish Academy of Music.

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

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition, 10,* 243–251.

Kuhl, P. K. (1988). Auditory perception and the evolution of speech. *Human Evolution, 3,* 19–43.

Langer, S. (1957). *Philosophy in a new key.* Cambridge, MA: Harvard University Press.

Laver, J. (1980). *The phonetic description of voice quality.* Cambridge: Cambridge University Press.

Lewis, M. M. (1951). *Infant speech.* London: Routledge & Kegan Paul.

Lieberman, P. (1984). *The biology and evolution of language.* Cambridge, MA: Harvard University Press.

Lieberman, P., & Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *Journal of the Acoustical Society of America, 34,* 922–927.

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

Malsheen, B. J. (1980). Two hypotheses for phonetic clarification in the speech of mothers to children. In G. H. Yeni-Komshian, J. F. Kavanaugh, & C. A. Ferguson (Eds.), *Child Phonology.* Vol. 2: *Perception* (pp. 173–184). New York: Academic Press.

Marler, P. (1990). Innate learning preferences: Signals for communication. *Developmental Psychobiology, 23,* 557–568.

Mayer, N. K., & Tronick, E. Z. (1985). Mothers' turn-giving signals and infant turning-taking in mother–infant interaction. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 199–216). Norwood, NJ: Ablex.

McCosker, S. S. (1974). *The lullabies of the San Blas Cuna Indians of Panama.* Gothenburg: Gothenburg Ethnographical Museum.

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

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19,* 498–550.

Morikawa, H., Shand, N., & Kosawa, Y. (1988). Maternal speech to prelingual infants in Japan and the United States: Relationships among functions, forms and referent. *Journal of Child Language, 15,* 237–256.

Morrongiello, B. A. (1988). The development of auditory pattern perception skills. *Advances in Infancy Research, 5,* 135–172.

Murray, A. N., Johnson, J., & Peters, J. (1990). Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language, 17,* 511–525.

Murray, L., & Trevarthan, C. (1986). The infant's role in mother–infant communications. *Journal of Child Language, 13,* 15–29.

Nettl, B. (1956). *Music in primitive culture.* Cambridge, MA: Harvard University Press.

Newman, R. A., & Emanuel, F. W. (1991). Pitch effects of vowel roughness and spectral noise for subjects in four musical voice classifications. *Journal of Speech and Hearing Research, 34,* 753–760.

Newport, E., Gleitman, L., & Gleitman, H. (1977). Mother, I'd rather do it myself: Some effects and noneffects of Motherese. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109–149). Cambridge: Cambridge University Press.

Papoušek, H., & Papoušek, M. (1984). Learning and cognition in the everyday life of human infants. *Advances in the Study of Behavior, 14,* 127–163.

Papoušek, M., Bornstein, H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development, 13,* 539–545.

Papoušek, M., & Papoušek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. *Advances in Infancy Research, 1,* 163–224.

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.

Papoušek, M., Papoušek, H., & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their three-month-old infants. *Journal of Psycholinguistic Research, 16,* 306–319.

Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development, 14,* 415–440.

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development, 15,* 325–345.

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

Rheingold, H., & Adams, J. L. (1980). The significance of speech to newborns. *Developmental Psychology, 16,* 397–403.

Rosenthal, M. K. (1982). Vocal dialogues in the neonatal period. *Developmental Psychology, 18,* 17–21.

Sachs, J. (1977). The adaptive significance of linguistic input to prelinguistic infants. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 51–61). Cambridge: Cambridge University Press.

Sachs, J., & Devin, J. (1976). Young children's use of age-appropriate speech styles in social interaction and role-playing. *Journal of Child Language, 3,* 81–98.

Sakata, H. L. (1987). Hazara women in Afghanistan: Innovators and preservers of a musical tradition. In E. Koskoff (Ed.), *Women and music in cross-cultural perspective* (pp. 85–95). Westport, CT: Greenwood Press.

Sands, K. M., & Sekaquaptewa, E. (1978). Four Hopi lullabies: A study in method and meaning. *American Indian Quarterly, 4,* 195–210.

Scherer, K. R. (1974). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S. Weitz (Ed.), *Non-verbal communication* (pp. 105–114). New York: Oxford University Press.

Scherer, K. R. (1981). Speech and emotional states. In J. K. Darby, Jr. (Ed.), *Speech evaluation in psychiatry* (pp. 189–220). New York: Grune & Stratton.

Scherer, K. R., Koivumaki, J., & Rosenthal, R. (1972). Minimal cues in the vocal communication of affect: Judging emotions from content-masked speech. *Journal of Psycholinguistic Research, 1,* 269–285.

Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion, 1,* 331–346.

Schneider, B. A., & Trehub, S. E. (1992). Sources of developmental change in auditory sensitivity. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 3–46). Washington, DC: American Psychological Association.

Schoenberg, A. (1975). *Style and idea* (L. Black, trans.). London: Faber & Faber.

Sherrod, K. B., Friedman, S., Crawley, S., Drake, D., & Devieux, J. (1977). Maternal language to prelinguistic infants: Syntactic aspects. *Child Development, 48,* 1662–1665.

Shute, B., & Wheldall, K. (1989). Pitch alterations in British motherese: Some preliminary acoustic data. *Journal of Child Language, 16,* 503–512.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language, 4,* 1–22.

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.

Stevens, K. N., & Keyser, S. J. (1989). Primary features and their enhancement in consonants. *Language, 65,* 81–106.

Stevenson, M. B., VerHoeve, J. N., Roach, M. A., & Levitt, L. A. (1986). The beginning of conversation: Early patterns of mother–infant vocal responsiveness. *Infant Behavior and Development, 9,* 423–440.

Stross, B. (1977). Tzeltal conceptions of power. In R. D. Fogelson & R. N. Adams (Eds.), *The anthropology of power: Ethnographic studies from Asia, Oceania, and the New World* (pp. 271–285). New York: Academic Press.

Sundberg, J. (1973). The source spectrum in professional singing. *Folia Phoniatrica, 25,* 71–90.

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.

Toda, S., Fogel, A., & Kawai, M. (1990). Maternal speech to three-month-old infants in the United States and Japan. *Journal of Child Language, 17,* 279–294.

Trainor, L. J. (1991). *The origins of musical pattern perception: A comparison of infants' and adults' processing of melody.* Unpublished doctoral dissertation, University of Toronto, Toronto.

Trainor, L. J., & Trehub, S. E. (1992a). A comparison of infants' and adults' sensitivity to Western tonal structure. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 394–402.

Trainor, L. J., & Trehub, S. E. (1992b). The development of referential meaning in music. *Music Perception, 9,* 455–470.

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

Trehub, S. E. (1985). Auditory pattern perception in infancy. In S. E. Trehub & B. A. Schneider (Eds.), *Auditory development in infancy* (pp. 183–195). New York: Plenum.

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., Endman, M., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology, 49,* 300–313.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm. Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology, 43,* 217–229.

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,* 5–19.

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. (1993). Listening strategies in infancy: The roots of music and language development. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: Cognitive perspectives on human audition* (pp. 278–327). London: Oxford University Press.

Trehub, S. E., & Unyk, A. M. (1992). Music prototypes in developmental perspective. *Psychomusicology, 10,* 31–45.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (in press-a). Adults identify infant-directed music across cultures. *Infant Behavior and Development.*

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (in press-b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development.*

Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music, 20,* 15–28.

Walker, R. (1990). *Musical beliefs: Psychoacoustic, mythical, and educational perspectives.* New York: Teacher's College Press.

Watkins, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception & Psychophysics, 37,* 275–288.

Watson-Gegeo, K. A., & Gegeo, D. W. (1986). Calling out and repeating routines in Kwara'ae children's language socialization. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (pp. 17–50). Cambridge: Cambridge University Press.

Weeks, T. (1971). Speech registers in young children. *Child Development, 42,* 1119–1131.

Werker, J. F. (1991). The ontogeny of speech perception. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 91–110). 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.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustic correlates. *Journal of the Acoustical Society of America, 52,* 1238–1250.

Wolff, P. H. (1963). Observations on the early development of smiling. In B. M. Foss (Ed.), *Determinants of infant behavior II* (pp. 113–134). London: Methuen.