The Development of Attention: Research and Theory
James T. Enns (Editor)

Chapter 5

Rules for Listening in Infancy

Sandra E. Trehub and Laurel J. Trainor

University of Toronto

In the present review, we outline a set of principles governing infants' deployment of attention to auditory events. Infants initially look in the direction of sounding objects, later reaching for the objects. They find some sound qualities highly salient, such as female voices in general and the mother's voice in particular. Infants selectively attend to the pitch contours and rhythms of animated speech and musical sequences. They encode finer details of some musical sequences, notably those typical of their culture. Moreover, simplicity or familiarity of the sequences and greater maturity of the infant lead to more comprehensive auditory processing. Finally, we identify a number of parallels between infants' processing of speech and music, and propose directions for future research.

Recent research has negated the view of the auditory world of the infant as undifferentiated. Rather, infants perceive structure in the sounds they hear (Morrongiello, 1988b; Trehub, 1985, 1987, 1989, 1990), attending to features that facilitate the localization, organization, and categorization of such patterns. Our goal, in the present review, is to derive a set of principles or heuristics governing infants' deployment of attention to auditory events. Our primary concern is not with the single decontextualized sounds of the typical laboratory experiment but rather with the rich and varied sound sequences that are prominent in the environment. We attempt, moreover, to relate the proposed principles to ecological considerations. In other words, how do such listening strategies facilitate the extraction of useful information about the world?

Rule 1: *At birth, turn toward the general direction of a sound source.* **Developmental Addendum:** *Increase the flexibility, rapidity, and accuracy of the response by 4-5 months of age, reaching as well as looking toward sounding objects.*

Head turning in the direction of off-centered sounds is demonstrable in the early days of life (e.g., Muir & Field, 1979), implying that visual and auditory attention are directed to informative areas of the environment. Despite the reliability of the response, definitive evidence of its occurrence in the neonatal period did not appear until 1979. One factor that tended to obscure the response is its long latency, typically reported as 8-11s following stimulus onset, which includes 5-7s for initiating a turn and another 3-4s for completing it (see Muir, Clifton, & Clarkson, 1989, pp. 202-203 for details). In situations with stimulus durations less than 5s, a response might be initiated well after stimulus offset (Clarkson, Clifton, & Morrongiello, 1985), thereby escaping the scrutiny of the most careful observer. Other potentially problematic factors are the presence of competing visual stimulation (Muir & Clifton, 1985), habituation of the response (Zelazo, Brody, & Chaika, 1984), and relatively intense sounds, which can induce head and eye movements away from the sound source (Butterworth & Castillo, 1976).

Sound localization in the horizontal plane exhibits a peculiar developmental timetable. Although most newborns turn reliably to sound, the response tends to disappear at around 1-2 months of age, reappearing at around 3-4 months of age (Field, Muir, Pilon, Sinclair, & Dodwell, 1980). The influence of maturational factors can be seen clearly with infants born prematurely; in such cases, the developmental schedule corresponds to time from conception rather than birth (Muir, 1985). The prevailing wisdom regarding the developmental progression is that the response is initially under subcortical control and reflexive in character; its disappearance is linked to reflex inhibition associated with the transition to cortical processing; and its reappearance to the full assumption of cortical control (Muir & Clifton, 1985; Muir et al., 1989).

Although it is difficult to elicit a head turn to off-centered sounds in 2-month old infants, it is somewhat less difficult when the stimulus is highly salient. For example, a female voice more readily provokes directional head turns (Field et al., 1980) than does the rattling of popcorn kernels (and acoustic analogues) found in typical investigations of infant sound localization. Clearly, optimal stimulus parameters for inducing a head turn response remain to be determined. There is evidence, however, for greater newborn responsiveness to higher than to lower frequency sounds (Morrongiello & Clifton, 1984). Whether this derives from differential sensitivity to high and low frequencies (Trehub, Schneider, & Endman, 1980) or from differential detectability of the disparate localization cues for high-frequency (interaural amplitude differences) and low-frequency (interaural time differences) signals remains unresolved.

In any case, reappearance of the sound localization response, presumably under cortical control, is marked by a number of changes, including a dramatic drop in the latency of responding (1.5s by 5 months of age), the appearance of active visual search at the locus of sound, and the robust quality of the response (Muir et al., 1989), obviating the need for special stimuli and procedures. Although turning to off-centered sounds is characteristic of infants from 4-5 months of age, localization precision in the horizontal plane is still imperfect. The smallest detectable shift in sound (i.e., minimum audible angle) is about 12°-19° for infants 6-7 months of age and about 4° for 18-month olds (Ashmead, Clifton, & Perris, 1987; Morrongiello, 1988a), compared to 1-2° for adults.

Localization in the vertical plane, which is dependent upon spectral differences resulting from the diffraction of sound waves by the pinna (Hebrank & Wright, 1974), exhibits a somewhat different developmental course. In the context of age-related changes in the size or shape of the pinna, an extended period of learning or calibration could be expected (Morrongiello & Rocca, 1987). In fact, newborns are relatively poor at localizing such sounds (Muir, 1985), but 6-months olds detect sounds displaced 15° vertically and 18-month olds detect 4°, which is comparable to adult performance (Morrongiello & Rocca, 1987). Just as high-frequency components facilitate the localization of signals in the horizontal plane, so do they influence vertical localization, this influence continuing throughout infancy (Morrongiello, 1987).

By 4-5 months of age, infants not only look in the direction of sound, they also reach in that direction (Stack, Muir, Sherriff, & Roman, 1989; Wishart, Bower, & Dunkfield, 1978). Moreover, prior visual exposure to sounding objects increases the accuracy of reaching for such objects in darkness, at least for 7-month olds (Perris & Clifton, 1988).

From an ecological perspective, early orientation of receptors toward the locus of sound stimulation maximizes information about the environment. The relevant mechanisms seem to come under voluntary control by about 4-5 months of age, at which time infants visually search and reach toward sounding objects. Well before then, however, nature provides reflexive mechanisms for orienting toward off-centered sounds.

Rule 2: ***Attend to female voices, particularly the mother's voice. Attend selectively to infant-directed speech.***

One of the most remarkable human capabilities is the recognition of specific voices across variations in speaking rate, intonation, and environmental circumstances that alter spectral and temporal information (e.g., room acoustics, telephone transmission). The precise cues underlying voice recognition are unknown, but the presumption is that we use some combination of fundamental frequency, timbre (i.e., sound quality), temporal patterning, and other prosodic cues, with different cues figuring more prominently in different voices (Abberton & Fourcin, 1978; Van Lanker, Kreiman, & Emmorey, 1985).

In the early months of life, when visual resolution is poor (Banks & Bennett, 1988), audition could provide a more efficient means of gaining information about the environment. Voice recognition, in particular, could foster the infant-caretaker bond, thereby increasing the infant's exposure to cognitive as well as social stimulation. Unfortunately, however, there has been relatively little research in this domain and the findings are somewhat difficult to interpret.

There are claims that newborns alter their pattern of nonnutritive sucking to maximize the availability of their mother's voice compared to that of a female stranger, or a female compared to a male voice (DeCasper & Fifer, 1980; Fifer & Moon, 1988). Previously, maternal voice recognition had been reported for infants one month or older (e.g., Mehler, Bertoncini, Barrière, & Jassik-Gerschenfeld, 1978). By contrast, the search for newborn preference for the father's voice over that of a male stranger has been unsuccessful (DeCasper & Prescott, 1984), leading some investigators to argue that the maternal voice preference arises from prenatal rather than postnatal exposure. How might this be accomplished?

For obvious reasons, it is difficult to describe the fetal auditory environment. There are suggestions, however, of substantial attenuation of frequencies above 1000 Hz and masking of very low frequencies (including father's voice) by the mother's vascular and digestive sounds (Querleu & Renard, 1981), so that the maternal voice may be the most prominent auditory signal in utero. In contrast to other voices, the mother's voice is transmitted through her bones as well as through the air and abdominal wall. Nevertheless, its effective intensity remains unclear because of uncertainty regarding the ambient intrauterine noise level (Querleu, Renard, Versyp, Paris-Delrue, & Crepin, 1988). Moreover, the fluid amniotic environment precludes typical air conduction through the external and middle ear, so that it is difficult to imagine what the fetus actually hears (presumably via bone conduction). The contention, however, is that the fetus hears the low-frequency components of maternal speech and as well as maternal cardiovascular sounds (DeCasper & Sigafoos, 1983).

Among the speech cues potentially available in utero are prosodic features, which include patterns of intonation, stress, and timing. Perhaps such cues provide a potential basis for newborns' reported preference for their mother's voice (DeCasper & Fifer, 1980). Indeed, prosody cues maternal voice recognition in 1-month old infants, who alter their nonnutritive sucking to hear their mother (rather than a stranger) speaking in the *babytalk* or *motherese* register but not in an atypical monotone (Mehler et al., 1978). What has not been evaluated is whether they prefer their mother's voice over that of a stranger in the typical adult register.

The prominent prosodic features of infant-directed speech include heightened pitch, increased pitch range, greater rhythmicity, slower tempo, briefer utterances, and simpler pitch contours compared to adult-directed speech (Fernald & Simon, 1984; Papousek & Papousek, 1981; Stern, Spieker, Barnett, & MacKain, 1983). Infants respond to such speech with smiling (Mayer & Tronick, 1985; Wolff, 1987), vocalization (Stevenson, Ver Hoeve, Roach, & Leavitt, 1986), and heightened affect (Werker & McLeod, 1989). Moreover, given a choice of infant-directed or adult-directed speech by turning to one loudspeaker or another, 4-month olds select infant-directed speech (Fernald, 1985), a selection based primarily on the pitch patterning of such speech as opposed to its temporal or amplitude patterning (Fernald & Kuhl, 1987). There is evidence, however, that 1-month olds prefer the full complement of spectral and temporal cues over pitch patterning alone (Cooper & Aslin, 1989).

How might the salience of infant-directed speech in the postnatal period be linked to prenatal exposure? To the extent that the fetus is exposed to maternal speech, such speech would be adult-directed and, as a result, lacking in infant-directed prosody. Prenatal exposure, then, should result in a preference for adult-directed over infant-directed speech, but there is no such evidence in newborns and the very opposite occurs in older infants (e.g., Fernald, 1985). In fact, an initial preference for adult-directed speech would be counter-productive in heightening attention to speech addressed to others at the expense of speech intended for infants themselves. Thus, prenatal exposure as the factor underlying postnatal voice recognition (Fifer & Moon, 1988) requires stronger evidence than that currently available. Such evidence could include the following: common prosodic elements in infant-directed and adult-directed speech; the availability of such elements in utero; and common timbral (voice quality) cues in the pre- and postnatal period. Similarly, the claim that the reinforcing efficacy of female voices in general and the mother's voice in particular stems from their likeness to familiar maternal sounds in utero (DeCasper & Spence, 1986) needs further substantiation. In fact, when unfamiliar female speech sounds are low-pass filtered (i.e., low frequencies remaining intact) to mimic *womb-like* speech, newborns prefer the unfiltered version (Spence & DeCasper, 1987), as do 1-month olds (Cooper & Aslin, 1989).

An alternative, albeit speculative, interpretation is that the preference for female over male speech stems from an inborn preference for voices with higher fundamental frequency, approaching the infant's own pitch range. This would account not only for the salience of female compared to male voices but also for the preference for infant-directed over adult-directed speech. In fact, infant-directed speech involves a 3-4 semitone increase in pitch (Fernald & Simon, 1984), resulting in female speech in the infant's vocal production range. Indeed, imitation of maternal intonation contours is evident as early as 6 weeks of age (Lieberman, Ryalls, & Robson, 1982, in Lieberman, 1984) and imitation of sung pitches by 3 to 6 months of age (Kessen, Levine, & Wendrich, 1979).

An associated speculation is that human infants, like those of other species (Alberts, 1981), quickly learn to differentiate highly salient stimuli from one another, particularly those associated with aspects of maternal care. For example, human neonates respond preferentially to odors associated with their lactating mother's breast milk (Macfarlane, 1975) and armpits (Cernoch & Porter, 1985). Is it surprising, then, that they also respond preferentially to their mother's voice? Although odors associated with lactation are attractive even to bottle-fed infants (Makin & Porter, 1989), breast-feeding experience is necessary for recognizing the mother's unique olfactory signature (Cernoch & Porter, 1985; Macfarlane, 1975). In analogous fashion, female voices may be inherently attractive to infants, but minimal exposure may be necessary for recognition of the mother's unique vocal signature. In any case, recognition of the mother's voice after mere hours of unattenuated and undistorted exposure remains an impressive accomplishment.

Rule 3: *Attend to the intonation contour of infant-directed speech.*

As noted, infants selectively listen to speech intended for them, paying particular attention to its intonation or pitch configuration (Fernald & Kuhl, 1987; Papousek & Papousek, 1981). There are suggestions, moreover, that the prosodic domain may involve unusually rapid learning. For example, 4-day old infants prefer the prosody of their native language over that of a foreign language (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988). Mothers reinforce this prosodic focus by delivering a limited number of simple pitch contours that they repeat frequently with altered lexical or segmental content (Fernald & Simon, 1984; Stern, Spieker, & MacKain, 1982).

Such prosodic modifications occur in numerous tonal and nontonal languages (Fernald, Taeschner, Dunn, Papousek, de Boysson-Bardies, & Fukui, 1989; Grieser & Kuhl, 1988; Papousek & Papousek, in press), increasing the likelihood that they are universal, if not identical in all respects (Fernald, in press; Fernald et al., 1989). Specific contours serve particular attentional and arousal regulatory functions, with rising contours used to capture infants' attention (Ferrier, 1985; Papousek & Papousek, 1981), bell-shaped contours (rise-fall or fall-rise) to maintain attention (Stern et al., 1982), and falling contours to comfort or soothe (Papousek, Papousek, & Bornstein, 1985). Recent evidence indicates, moreover, that the contours of maternal speech carry a range of communicative intentions that are transparent to adults (Fernald, 1989) and even to infants (Fernald, in press). For example, adults can relate different prosodic contours to appropriate interactional contexts (e.g., approval, prohibition, comfort) and they do so more accurately for infant-directed than for adult-directed speech (Fernald, 1989). Similarly, 4-month olds exhibit positive affect to infant-directed vocalizations of approval and negative affect to prohibition vocalizations (in unfamiliar as well as familiar languages), implying that pitch contour or intonation is a key distinguishing feature (Fernald, in press).

The very contours that characterize infant-directed speech also predominate in the infant's own vocalizations (Delack & Fowlow, 1978; Papousek & Papousek, 1981). Despite their inability to match maternal utterance duration in the early weeks of life, infants nevertheless imitate maternal contours in their much briefer utterances (Lieberman, Ryalls, & Robson, 1982, cited in Lieberman, 1984). Such an achievement would seem to depend upon a global representation of pitch contour.

Maternal prosody has important functions that go beyond the modulation of infant attention and affect. There are indications that such prosody is replete with cues to the structure of language (e.g., Morgan, Meier, & Newport, 1987). For example, rising pitch contours and pause placement serve as reliable cues to clause boundaries, with infants being sensitive to such cues in unfamiliar as well as familiar languages (Jusczyk, 1989; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989). In this way, prosodic highlighting of key linguistic units may facilitate language acquisition (Kemler Nelson et al., 1989). Indeed, infants produce native-like prosodic frames well before their first words (de Boysson-Bardies, Sagart, & Durand, 1984; Crystal, 1973).

Rule 4: *Attend to the melodic contour of nonspeech sequences.*

Is infants' strategy of attending to contour applied exclusively to infant-directed speech or is it also used with other auditory (nonspeech) patterns? Adults' perceptual strategies with nonspeech sequences provide a useful backdrop to the consideration of infants' pattern perception skills.

It has long been clear that melody recognition is independent of specific notes or pitch levels, depending instead on relations among component notes. When listening to familiar melodies, adults attend to the pattern of successive *intervals* (Attneave & Olson, 1971; Bartlett & Dowling, 1980), that is, the precise relation or distance between adjacent notes (designated in semitones or pitch ratios). It follows, then, that transposing a melody by changing all of its notes but maintaining the pattern of intervals would preserve the essence or identity of that melody. Further, it is not the case that we retain the exact pitches, simply ignoring them for conventional purposes of melody identification. Rather, we experience considerable difficulty recalling or recognizing the exact pitches of familiar tunes, even those heard only at one pitch level (Attneave & Olson, 1971). What this implies is that our representation of familiar melodies includes little, if any, information about component pitches (although perhaps some information about general pitch range), with exact interval information predominating. With unfamiliar melodies, our representation tends to exclude interval as well as exact pitch information, being restricted largely to configurational information about successive directional changes in pitch or what is termed *melodic contour* (Dowling, 1978). Contour, in this context, refers only to directional aspects of pitch changes (rising, falling, staying the same), not their extent (Dowling & Harwood, 1986).

If infants had limited exposure to a melody, what features would they retain for comparison with changed or novel melodies? If we presume that adults' relational processing strategy derives from experience or conventions associated with language or music, then infants would hardly be in a position to adopt such a strategy. Instead, they might retain the pitch sequence in exact form, limited only by their memory capacity. This would be somewhat analogous to a tape recording of limited duration, yielding a faithful copy of the melody, if very brief, and the initial or final portion of the melody, if longer. If infants opted, instead, for a summary description of the melody, in line with their presumed representation of infant-directed speech (Fernald & Kuhl, 1987), then they might retain its contour. If, however, the melody was a highly salient stimulus, like that of their mother's voice (DeCasper & Fifer, 1980; Mehler et al., 1978), then infants might extract and retain more detailed information that uniquely defined the melody. The latter case might imply some biological disposition for rapid learning.

The relational processing of auditory information confers obvious advantages by providing listeners with structured or integrated information about the environment. It would not be surprising, then, for such auditory pattern processing abilities to be available in early life, as they are in the visual domain (Bornstein & Krinsky, 1985; Humphrey & Humphrey, 1989).

Chang and Trehub (1977a) first explored the possibility of relational coding of auditory sequences in infancy. In a habituation-dishabituation design, they familiarized infants with a sequence of six sinusoidal tones and then presented them with a transposition of the original (relational information preserved) or with a control pattern consisting of the notes of the transposition in different order (relational information changed). Response recovery was evident to the control pattern, indicating that infants differentiated it from the original. By contrast, there was no response recovery to the transposition, implying that common relational information led to its perception as identical or substantially similar to the original.

Although infants had encoded relational information from the tone sequences, the specific nature of that relational information remained unclear. Did infants use a contour processing strategy, as adults do with unfamiliar melodies (Bartlett & Dowling, 1980), or did they use an interval-processing strategy, as adults do with familiar or highly overlearned melodies (Dowling & Fujitani, 1971)? Over the past several years, we have been pursuing questions such as these with a technique that capitalizes on infants' inclination to turn toward interesting auditory events (see Eilers, Wilson, & Moore, 1977; Kuhl, 1985), adapting it for use with sound sequences (see Trehub, Bull, & Thorpe, 1984) rather than single sounds. In short, we familiarize infants (typically 6 to 11 months) with a particular sequence of sinusoidal tones and then test for their response to subtle or substantial deviations from that sequence. In this way, we can gather information about aspects of a pattern that underlie its perception as familiar or novel. We use sinusoidal tones to minimize potential cues from overlapping harmonics of various component tones. Thus, although individual tones are harmonically impoverished or unnatural, the sequences of such tones are structurally rich.

All testing is conducted in a sound-attenuating booth according to the following procedure. We present a repeating pattern over a loudspeaker to one side and periodically substitute a pattern that is altered in some respects (see Trehub et al., 1984; Trehub, Thorpe, & Morrongiello, 1987 for details). The tester and attending parent wear headphones with masking patterns so that they remain unaware of stimulus changes presented to the infant. The tester presses one button to indicate when the infant is attentive and looking directly ahead (i.e., ready for a trial) and another button when the infant turns 45° to the loudspeaker. The computer monitors such turns and the occurrence of targeted stimulus changes, delivering visual reinforcement when a turn occurs within 4 seconds of a change.

We determine, first, that infants can perform the task of turning to a sound change by requiring them to meet a training criterion of 4 successive correct responses to a salient change within 20 trials. In the subsequent test phase, we present approximately 30 test trials, half involving changes in the melody, half involving no change (randomly ordered). Since spontaneous turns toward the loudspeaker are likely to occur, no-change or control trials are necessary for interpreting performance on the change trials. Turning significantly more on change than on no-change trials indicates that infants detect and are responsive to the change in question. Interpreting the absence of a difference is more difficult because this can arise from failure to detect the difference or from infants' perception of the comparison pattern as substantially similar to the original.

In one study (Trehub et al., 1984), infants were tested for their discrimination of a number of changes to a six-tone melody, among them, transpositions (different notes, same intervals and contour), contour-preserving changes (different notes and intervals, same contour), and contour-violating changes (same notes reordered, resulting in different intervals and contour). Sample melodies are illustrated in Figure 1. For adult listeners, the transposed melodies would be most similar to the original and those with different contour least similar. When the retention interval was very brief (Experiment 1: 800 ms between standard and comparison sequences), infants responded to all changes, although their performance was most accurate for the contour changes. When the task was made more difficult by lengthening the retention interval (Experiment 2: 2.6 seconds between standard and comparison sequences) and filling it with distractor sequences, infants no longer responded to the transpositions and same-contour sequences but still responded to the contour changes. In other words, precise pitch and interval information seemed to be available immediately but to decay rapidly, leaving contour information intact. In subsequent research, it became apparent that infants could detect contour changes even when a single note of a six-tone sequence brought about a directional change in pitch (Trehub, Thorpe, & Morrongiello, 1985). What also emerged was that infants encode information about frequency range (Trehub et al., 1984, 1985), treating comparison melodies as familiar if they have the same contour and approximate frequency range as the standard melody and as novel if either the contour or range differs.

Although it was clear that infants could detect differences between melodies that differed in contour, it was less clear whether they perceived similarities between patterns with common contour. In other words, was it possible that infants could notice the differences but nevertheless respond equivalently? In the context of a single repeating standard and comparison melody (e.g., Trehub et al., 1984), this question could not be resolved. One means of doing so was to familiarize infants with sets of discriminable melodies that shared a common contour, requiring them to differentiate that set from another set of melodies with contrasting contour (Ferland & Mendelson, 1989; Trehub et al., 1987). The task is illustrated in Figure 2. Infants responded to such contour changes in the context of changes in pitch, interval size, and waveform, indicating that they could categorize melodic sequences on the basis of contour.

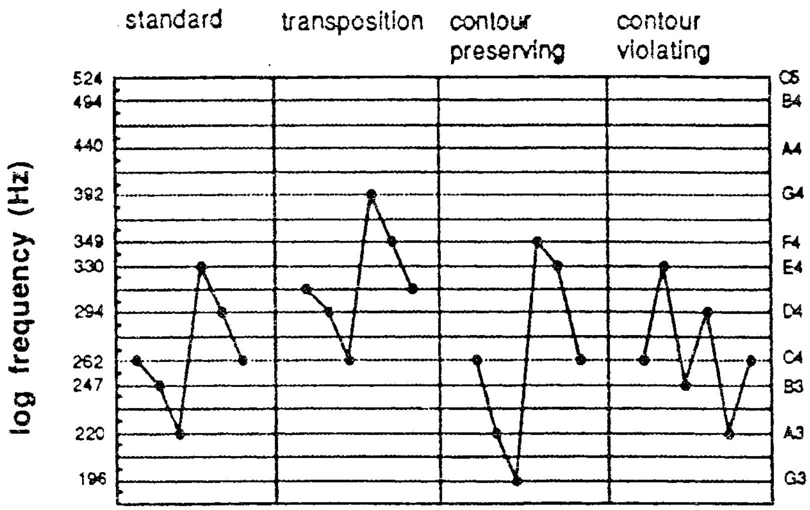
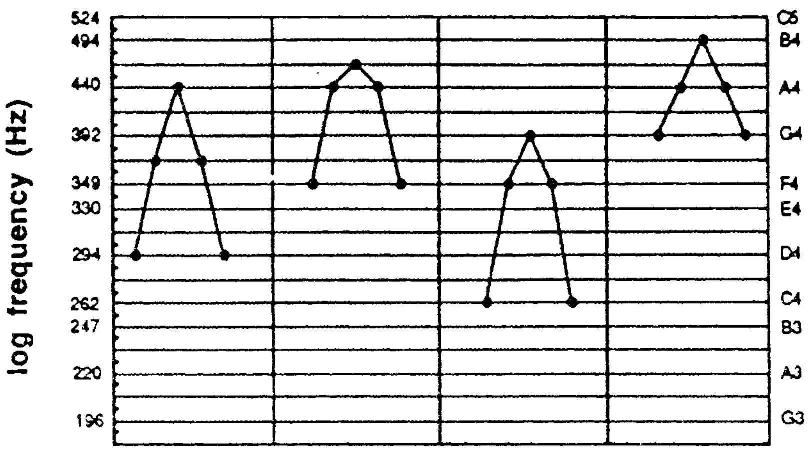


Figure 1. Sample melodies from Trehub et al. (1984): standard melody in C major, transposition to E^b major, contour-preserving transformation, and contour-violating transformation. Successive horizontal lines represent semitones. Letters represent note names and associated numbers denote the relevant octave (e.g., C4 is middle C and C5 is one octave higher).

Rule 5: *Attend to the temporal (rhythmic) structure of auditory sequences and group the elements of such sequences.*

As we have seen, the pattern of pitch directional changes in an auditory sequence generates a contour that is important in infants' as well as adults' representation of melodies. For adults, at least, the pattern of relative onset times of notes also generates a unique temporal or rhythmic identity. Nevertheless, temporal aspects of auditory sequences have received much less experimental attention than have melodic aspects (Dowling & Harwood, 1986) despite their role in speech comprehension (Eimas & Miller, 1980; Martin, 1972) and memory (Dowling, 1973; Huttenlocher & Burke, 1976).



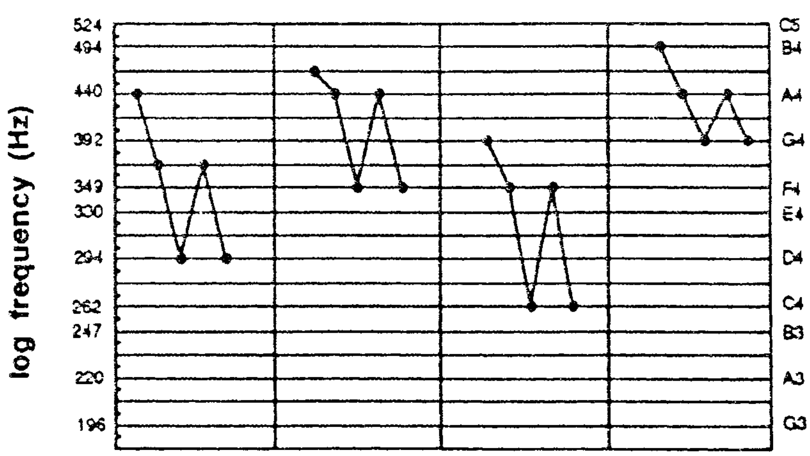


Figure 2. Sample melodies from Trehub et al. (1987). Upper panel: set of contour-preserving (variable interval) melodies. Lower panel: contrasting set with changed contour. Successive horizontal lines represent note names and associated numbers denote the relevant octave.

We do know that infants can discriminate between auditory sequences with contrasting temporal structure (Chang & Trehub, 1977b; Demany, McKenzie & Vurpillot, 1977; Morrongiello & Trehub, 1987) and that they can detect temporal correspondences between auditory and visual patterns (Bahrick, 1983; Mendelson & Ferland, 1982; Spelke, Born & Chu, 1983) but the role of relational processing in this domain has only emerged recently.

Trehub and Thorpe (1989) explored infants' ability to discriminate patterns with contrasting temporal structure (X XX vs. XX X or XX XX vs. XXX X) in the context of tempo (speed) and pitch variations. In other words, repetitions of the patterns involved varying pitch levels and tone durations but preserved the relative timing of elements. Infants succeeded in differentiating such tone sequences and also in categorizing them on the basis of temporal structure. These findings reveal infants' disposition to apply relational processing strategies to the temporal domain. What remains unclear is the extent to which comparable strategies are used with speech input. Although infant-directed speech is noted for its rhythmicity (Beebe, Feldstein, Jaffe, Mays, & Alson, 1985; Fernald & Simon, 1984), the absence of detailed temporal analyses leaves several questions unanswered, including the possibility of maternal timing universals, the impact of specific rhythmic patterns on infant affect, and potential temporal cues to maternal voice recognition. In any case, it is of interest that mothers spontaneously adopt a rhythmic presentation strategy, given its known facilitation of cognitive processing (Huttenlocher & Burke, 1976).

Another critical aspect of temporal processing concerns our propensity to *group* elements within auditory sequences. Generally, such groupings are promoted by the relative temporal proximity of elements (Chang & Trehub, 1977b; Trehub & Thorpe, 1989) or by their similarity (Demany, 1982; Thorpe & Trehub, 1989), but perceptual groups can also emerge in sequences with totally uniform elements (Fraisse, 1982). Indeed, the appearance of perceptual groups in the context of identical elements serves to emphasize the strength and prevalence of these grouping processes.

It is likely that grouping processes generate the perceptual chunks or units for further cognitive processing (Bregman, 1981) and it is possible, moreover, that such grouping processes are relatively immune to the listener's knowledge or experience (Fodor, 1983). These considerations increase the likelihood that grouping mechanisms would be operative in infancy, promoting adult-like interpretations of sensory input.

Grouping, with its attendant segregation of some elements from others, has certain perceptual consequences, such as creating the illusory sense of pauses between words (Studdert-Kennedy, 1975) or tone groups (Bolton, 1894; Thorpe, 1985). Another likely consequence of grouping is that changes that conserve the original grouping structure should be more difficult to detect than those that disrupt it. For example, temporal alterations of speech are relatively difficult to notice if the overall timing between accented syllables is preserved (Huggins, 1972).

Infants 6 to 9 months of age were presented with tone sequences characterized by the following structure: XXXOOO (Thorpe & Trehub, 1989; Thorpe, Trehub, Morrongiello, & Bull, 1988). The standard patterns consisted of six temporally equidistant tones; the first three were identical, contrasting with the last three (in pitch, waveform, or intensity), which were also identical to one another. The comparison patterns embodied an extended intertone interval or pause, either at the boundary between hypothesized groups (XXX OOO) or within a group (XXXO OO). Whereas the former comparison pattern (between-group pause) was consistent with the structure of the original pattern, the latter comparison (within-group pause) was not. The findings revealed that the structure-disrupting changes were more readily detectable than were the structure-conserving changes, implying that infants grouped the elements of these sequences on the basis of their similarity. In a somewhat similar vein, Krumhansl and Jusczyk (1990) inserted pauses *between* the phrases of Mozart minuets or *within* such phrases and presented infants with a listening choice. Infants as young as 4 1/2 months exhibited listening preferences for musical samples that preserved the appropriate musical phrase structure (i.e., between-phrase pauses). Since infants had not been familiarized with the original or intact sequences, as they were in the aforementioned studies (Thorpe & Trehub, 1989; Thorpe et al., 1988), this suggests that they are sensitive to the integrity of musical phrase structure.

Some parallels are evident in recent research on running speech. For example, infant-directed speech that is manipulated by the insertion of pauses *within* clauses disrupts infant attention more than pauses inserted *between* clauses, an effect that is absent in adult-directed speech (Kemler Nelson et al., 1989). This implies that cues to the grouping or segmentation of auditory sequences are prominent in infant-directed speech, as they are in music. Thus, there is unequivocal evidence of infants' temporal analysis and organization of both speech and tone sequences.

Rule 6: ***Encode good patterns with relative ease (i.e., lesser effort), assimilating greater detail from such patterns.***

Patterns with so-called *good form* or structure tend to be perceived and remembered more readily. But what is *good form*? In the visual domain, there have been numerous attempts to characterize it, either with reference to primitive coding properties of the visual system (Hoffman & Dodwell, 1985) or by stimulus descriptions that translate Gestalt concepts of figural *goodness* into information theoretic terms (Garner, 1970, 1974). Moreover, there has been considerable discussion of the importance of symmetry in visual perception (e.g., Hoffman &

Dodwell, 1985), including the proposal that symmetry is extracted at an early stage of processing, even prior to the identification of a display (Corballis & Beale, 1976).

Research emanating from such ideas has revealed that figural goodness has implications for infants' perception and categorization of visual patterns (Humphrey & Humphrey, 1989). For example, 3-month olds can categorize good patterns, 5-month olds, intermediate patterns, and 7-month olds, poor patterns, as defined by Garner's (1970) system (Younger & Gotlieb, 1988). Moreover, infants are sensitive to the bilateral symmetry of visual patterns (Bornstein & Krinsky, 1985), preferring symmetrical over asymmetrical patterns (Humphrey & Humphrey, 1989). Nevertheless, a definition of pattern *goodness* that applies to all known phenomena of visual perceptual organization remains elusive (Humphrey & Humphrey, 1989).

There has been considerably less progress in specifying the dimensions of *goodness* of auditory patterns. A number of researchers have attempted to describe the factors that promote the *temporal coherence* or connectedness of auditory sequences (e.g., Bregman, 1981). Clearly, patterns perceived as coherent would be better in some sense than those perceived as less coherent. Among the factors presumed to promote such coherence are the pitch proximity of successive sounds (Bregman & Campbell, 1971; Demany, 1982), gradual and smooth trajectory from one pitch level to another (Bregman & Dannenbring, 1973), simplicity of contour (Divenyi & Hirsh, 1974; McNally & Handel, 1977), and similar timbre of elements (McAdams & Bregman, 1979; Singh, 1987; Tougas & Bregman, 1985). Moreover, simple harmonic relations seem to promote the integration or fusion of simultaneous components (Moore, Glasberg, & Peters, 1986).

Infant-directed speech exemplifies some of these coherence-promoting properties, an obvious one being continuity of timbre (i.e., voice quality). Although larger pitch excursions are used in infant-directed than in adult-directed speech (presumably to attract attention), these excursions are linked by pitch glides (Fernald, 1984), which are effective in preserving coherence (Bregman & Dannenbring, 1973). Moreover, simple contours, frequently unidirectional (ascending or descending), are seen commonly, as noted earlier.

Another approach to the question of *goodness* involves the concept of prototypicality. Many categories of objects seem to be organized around a best example or prototype (Rosch & Mervis, 1975), with natural categories having universal prototypes (Mervis & Rosch, 1981). There have been suggestions that speech categories also have typical or *good* exemplars (e.g., Miller, 1977). Grieser and Kuhl (1989) exposed 6-month old infants to *good* or *poor* exemplars of the synthesized vowel [*i*] (as judged by adults) and tested for generalization to variants around the *good* or *poor* stimulus. Infants showed broader generalization around the *good* stimulus, implying that it functioned as a more stable category center.

When musical patterns are considered *good*, the frame of reference is also prototypicality or adherence to music theoretic principles. According to Jones (1981), musical prototypes or ideals summarize the conventions of a musical culture and involve harmonic, melodic, and rhythmic symmetries. Such prototypes are internalized on the basis of experience, creating expectancies that guide the listener's attention to musical sequences. These expectancies serve as a perceptual reference or backdrop against which unfolding melodies can be evaluated. It follows, then, that changes or deviations will be more attention-eliciting in the context of a prototypical standard (Jones, 1982). Indeed, there is considerable evidence in support of this position. For Western listeners, musical patterns that conform to Western scale structure are encoded in more detail, remembered more readily, and have a more stable representation or cognitive anchoring than those that are non-conforming (Cuddy, Cohen, & Mewhort, 1981; Krumhansl, Bharucha, & Kessler, 1982). Moreover, school-age children exhibit a preference for musical patterns that exemplify various aspects of Western tonal structure (Krumhansl & Keil, 1982). These effects are presumably attributable to musical exposure, although the extent of required exposure is unclear.

Familiarity with some details of Western tonal music may be necessary for clarifying the issues in question. The octave (pitch ratio of 2:1) is divided into 12 equal intervals or semitones that comprise the *chromatic scale* (notes 0, 1,...12, where notes 0 and 12 are one octave apart), a division that is repeated in successive higher and lower octaves. It is of interest, however, that Western musical compositions are not based on the equal-interval chromatic scale but on the unequal-interval *diatonic scale*, which involves a subset of 7 notes from the chromatic scale. (Other cultures also use unequal-interval scales, although they differ from ours in a number of respects.) For example, designating the *tonic* or reference note (the key of the melody) as note 0, the major scale consists of notes 0, 2, 4, 5, 7, 9, 11, 12. Thus, once the key of a melody is specified, the set of notes is also specified, including those occupying the same position (i.e., having the same note name) in other octaves. (Within an extended piece of music, occasional modulations to different keys are not uncommon, although the piece usually ends in the key in which it began. Ornamental notes may add to the interest of a piece by going outside the notes of the key.) Different notes within a key or scale have somewhat different roles. For example, the tonic note recurs frequently, confers a sense of stability and coherence, and commonly functions as the ending note. Moreover, the notes of the *major triad* (e.g., notes 0, 4, and 7 of the chromatic scale), consisting of *do mi sol* (C E G in the key of C: +4, +3 semitones), occur frequently in Western music (Roberts & Shaw, 1984; Simonton, 1984), the major triad being considered a prototype or primordial unit of tonal structure (Cuddy & Badertscher, 1987; Schenker, 1906/1954), with a very stable representation (Krumhansl et al., 1982).

Infants and preschool children were tested for their ability to discriminate a semitone change in the context of one of two five-tone melodies (Trehub, Cohen, Thorpe, & Morrongiello, 1986). One melody, C E G E C, had diatonic tones only, in particular, the tones of the major triad (see Figure 3, upper panel). The other melody, C E $G^{\#}$ E C (augmented triad), had the same contour but had one note, $G^{\#}$, that was nondiatonic or outside the scale in question (C major). Infants and children listened to exact repetitions (i.e., identical pitches) of the diatonic or nondiatonic melody and, periodically, to the comparison melody, which had a semitone change in any one of the five positions of the melody. If diatonic structure has functional priority for infants and preschool children, as it does for older children and adults, it should provide an enhanced context for detecting such subtle deviations. Infants and children detected semitone changes in the context of both musical sequences, indicating that they could encode the smallest intervals that are relevant to Western music. Only the preschoolers, however, showed superior performance in the diatonic context.

Although infants discriminated semitone differences that did not change the contour of the melody (i.e., interval differences), they did so under conditions of exact repetition of the standard and comparison melodies and very brief retention intervals (800 ms). These conditions are comparable to those of Trehub et al. (1984, Experiment 1), in which infants detected various changes that were not detectable under more difficult conditions such as longer retention intervals (e.g., Trehub et al., 1984, Experiment 2) or transposition of the standard and comparison sequences (Trehub et al., 1987). It is possible, then, that infants were simply detecting a change in the pitch level of one note rather than responding on the basis of relational pitch cues such as intervals. Transposing the melodies, as in Trehub et al. (1987), would eliminate absolute pitch cues and reveal whether infants could encode interval information.

In a further study (Cohen, Thorpe, & Trehub, 1987, Experiment 1), infants detected semitone changes in the context of transposed melodies based on the major (C E G E C) or minor (C E^{b} G E^{b} C) triad, the latter of which is also considered to be well structured (Roberts & Shaw, 1984). They performed significantly more poorly, however, when the background melody was less well structured (C E $G^{\#}$ E C) by virtue of one note being outside the key (Cohen et al., 1987, Experiment 2). An interesting asymmetry in performance emerged when the very same *good* and *poor* melodies served, in turn, as standard or comparison sequences. The *poor* comparison (C E $G^{\#}$ E C) was readily detected against the *good* standard (C E G E C) but the *good* comparison (C E G E C) was difficult to detect in relation to the *poor* standard (C E $G^{\#}$ E C). Similarly, adults find it easier to detect deviations from *good* (i.e., well structured) melodic (Bharucha, 1984) or rhythmic (Bharucha & Pryor, 1986) patterns than from *poor* (i.e., less well structured) patterns, even when the deviations from *poor* patterns result in *good* patterns. The notion is that a *good* standard promotes a relatively stable mental representation, which, in turn, facilitates comparisons with other stimuli. In short, it appears that infants can encode interval size, doing so most readily when the melodies exemplify *good* structure. Moreover, the observed asymmetry in performance may be related to the stability of the *good* sequence, as is the case for adults (Cuddy et al., 1981; Krumhansl et al., 1982) and possibly preschool children (Trehub et al., 1986).

Although infants had exhibited an ability to discriminate intervals in transposed contexts, the generality of this finding was unclear. All of the standard melodies had consisted of five tones (only three of which differed) and symmetrical configurations (e.g., C E G E C). Was interval processing in infancy limited to simple, symmetrical patterns with few notes? In adulthood, melodic symmetry enhances aesthetic preference (Balch, 1981) and memory (Dowling, 1972).

To extend the previous study of interval processing in infancy and clarify these issues, we presented infants with one of three five-note melodies in transposition (Trehub, Thorpe,& Trainor, in press) (see Figure 3, lower panel). As in Cohen et al. (1987), all patterns had a similar contour (rise-fall) and overall pitch range. By contrast, however, the patterns had five different component notes. One of the melodies, the only *good* one, had component notes consistent with a diatonic scale (B D G E C); it sounded like a typical Western tune. The second melody was *poor* or nondiatonic in that its notes (C F# B F C#) did not belong to any scale; it contained two so-called *dissonant* intervals (C F#and B F) and, moreover, it sounded unpleasant. The set of pitches of the third melody could not be placed in a chromatic or diatonic scale, two of its four intervals being smaller than a semitone. If infants failed to resolve the very small intervals in this context, then they would perceive only three rather than five different notes. The results were consistent with the previous findings. Infants were successful in detecting the semitone change in the *good* melody but were unsuccessful with either of the *poor* melodies. This indicates, at the very least, that neither symmetry nor a three-note set was essential for engaging an interval-processing strategy in infants. But what can account for infants' differential processing of *good* and *poor* melodies?

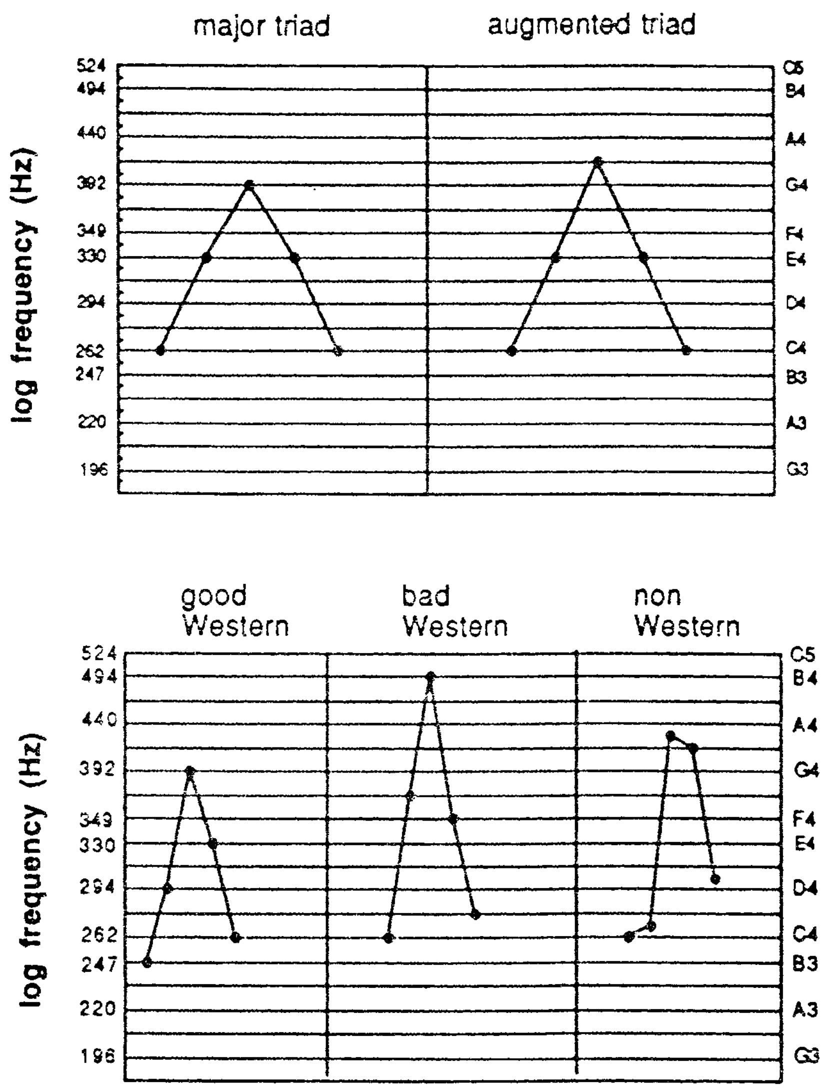


Figure 3. Upper panel: the *good* (major triad) melody and *poor* (augmented triad) melody used in Trehub et al. (1986). Lower panel: the *good* Western melody, *poor* Western melody, and non-Western melody used in Trehub et al. (in press). Successive horizontal lines represent note names and associated numbers denote the relevant octave.

The *good* melody was based on major triads, as was one of the *good* melodies in Cohen et al. (1987), the other one being based on the minor triad, another so-called *good* pattern (Roberts & Shaw, 1984). As noted earlier, the major triad is central to conceptions of Western tonal structure (Schenker 1906/1954) but is not known to be significant cross-culturally (Dowling & Harwood, 1986). It is interesting that the notes of the major triad embody very simple ratio relations (4:5:6) compared to the complex ratio relations of the *poor* melody in Cohen et al. (C E $G^{\#}$ E C; 16:20:25) and the more complex ratio relations of the *poor* melodies in Trehub et al. (in press). Perhaps infants' exposure to the small ratios of naturally occurring sounds is a contributing factor. There are suggestions that musical tuning systems evolved to maximize sets of notes related by small integer ratios (Terhardt, 1978; Watkins, 1985). Moreover, there are group-theoretical justifications for characterizing the major triad as an intrinsically *good* form. Balzano's (1982) mathematical derivation of diatonic scale structure generates the conclusion that the most common chords of the diatonic scale, including major and minor triads, form optimally compact spatial schemes. Thus, the views of various music and psychoacoustic theorists converge in their conception of the major triad as embodying special properties. Empirical research supports such a notion, indicating that the major triad is even more effective in establishing a sense of key and tonal structure than is the entire major scale (Cuddy & Badertscher, 1987).

Another feature shared by all of the *good* melodies presented to infants was the prominence of fifths (i.e., the interval formed by the outer notes of the major and minor triads). In studies of the relative significance of various notes of the diatonic and chromatic scale, the fifth note of the diatonic scale emerges as next in importance to the tonic note (Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979). In any case, if it is diatonic scale structure that underlies the superior performance with *good* melodies, then infants should perform poorly on *good* melodies from foreign cultures that use alternative scale structures. If, on the other hand, infants perform well on such prototypical foreign melodies, which presumably exemplify other kinds of *good form*, then *good form* in general rather than diatonic structure in particular would be implicated. These are testable hypotheses that are currently being evaluated in our laboratory.

If infants perform poorly on foreign prototype melodies, this would suggest that exposure to Western music, however limited and unsystematic, has resulted in the internalization of some aspects of Western scale structure or diatonicism. In fact, the frequency of occurrence of major triads in nursery or infant-directed songs (Cohen et al., 1987) is even greater than it is in adult-directed music (Simonton, 1984). Thus, to the extent that musical exposure does occur, one could count on the predominant exposure to major triadic relations. But is it reasonable to expect such musical exposure effects to be evident in the first year of life?

In the linguistic domain, effects of input language on phonetic perception can be seen by 10-12 months of age (Best & McRoberts, 1989; Werker & Lalonde, 1988; Werker & Tees, 1984), at which time infants experience difficulty with non-native speech contrasts that were discriminable at 6 months of age. Presumably, some reorganization of phonetic perception is related to the emergence of phonemic categories, which signal differences in meaning as well as sound (Best, McRoberts, & Sithole, 1988; Werker, in press).

Other effects of listening experience appear even earlier. For example, preference for the native language over a foreign language (Mehler et al., 1988) and for the mother's voice over other voices (DeCasper & Fifer, 1980; Mehler et al., 1978) are evident in the early days of life. In the context of such dramatic experiential effects, the extraction of prominent relational features from music some months later is not inconceivable.

Rule 7: *Proceed from global to local processing of auditory sequences.*

There are various ways in which infants are likely to proceed from global processing of musical or speech sequences to the processing of smaller events embedded therein. In the pitch domain, tunes can be considered in terms of a contour, a pattern of intervals, or a collection of pitches, in order of increasing specificity. In the temporal domain, patterns can be considered in terms of rhythmic structure, tempo, and note and silence durations. Similarly, the speech stream can be considered in terms of utterances, phrases, words, syllables, phonemes, and phonetic features.

Prelinguistic listeners, when confronted with a sequence that exceeds their immediate memory capacity (e.g., a tune or utterance), seem to attend to global features, extracting the contour and pitch range. When the stimulation or context poses fewer memorial demands, such as single sounds or fixed repeating sequences, infants analyze the material further, encoding, for example, the exact pitches or intervals of tunes (Trehub et al., 1984, 1985, 1986) or the phonetic segments of syllables (e.g., Kuhl, 1987; Trehub, 1979). In this regard, it is important to note that the typical phonetic perception experiment with infants involves single sounds rather than sound sequences.

Increasing familiarity with a sound pattern or age-related changes in memory capacity might make it possible for infants to simultaneously process global and local aspects of auditory materials. Thus, for example, infants extract the intervals as well as the contour of prototypical melodies (Cohen et al., 1987; Trehub et al., in press). It is possible that this reflects perceptual reorganization comparable to that observed in the phonetic domain (Werker & Lalonde, 1988; Werker & Tees, 1984). Just as infants begin by being sensitive to a broad range of speech contrasts and then narrowing this range of sensitivity by about 1 year of age (Best & McRoberts, 1989; Werker & Tees, 1984), so infants may be sensitive initially to foreign as well as native musical structures, with later facilitation for native structures reflected in their enhanced performance for prototypical Western tunes.

There is no better example of a familiar or prototypical auditory stimulus for infants than their mother's voice. To date, however, discussions of infants' responsiveness to mother's voice have focused primarily on its contour, without consideration of more detailed aspects of its pitch patterning. It is possible, then, that infants go beyond a contour processing strategy, encoding the precise extent of the mother's pitch excursions or intervals. This would provide them with a basis for recognizing their mother by her unique yet familiar tunes, which might also be presented in a personalized set of rhythms. Some suggestive evidence is provided by cross-cultural differences and mother-father differences in pitch excursions in infant-directed speech (Fernald et al., 1989). There is no information, however, on typical within-culture variations in maternal pitch excursions. (See Bettes, 1988 for the effects of maternal depression on such pitch excursions.)

Once maternal tunes successfully attract the infant's attention and are processed without difficulty, infants can proceed to analyze the words and sounds therein. In this way, musical aspects of the speech stream can guide the infant's entry into the linguistic domain. The prosodic aspects of speech continue to be informative to the language learner, marking clause and phrase boundaries (Kemler Nelson et al., 1989).

Summary and future directions

The foregoing rules or heuristics provide a preliminary description of the infant's auditory world. The infant first looks in the direction of sounding objects, later reaching for these objects. Some sound qualities are highly salient, such as female voices in general and the mother's voice in particular. Infants selectively attend to the pitch contours and rhythms of animated speech and musical sequences, perhaps identifying speakers and languages on these bases. They encode finer details of some musical patterns, notably those typical of their culture. Finally, simplicity or familiarity of the patterns and greater maturity of the infant lead to more comprehensive processing of auditory input.

The parallels between infants' processing of speech and musical sequences are intriguing and merit further exploration. Are musical and linguistic processes intertwined in early life because of shared auditory mechanisms, shared social-emotional functions, or both? Do primitive arousal-regulating and emotion-inducing aspects of music and music-like stimuli (e.g., infant-directed speech) motivate the initial steps of the language-learning process? Or is it simply the case that our emotional reactions to music are rooted in primitive affectional ties to our primary caretaker?

The musical quality of infant-directed speech suggests the possibility of even further parallels between speech and music. One musical form, the *lullaby*, has many features in common with infant-directed speech. The lullaby is also an intimate, aural communication between caregiver and infant, "a song sung in love to an audience of one" (Cass-Beggs & Cass-Beggs, 1969, p. 5). Infant-directed song, like infant-directed speech, seems to be characterized by higher overall pitch, wider pitch range, shorter phrases, and greater rhythmicity than adult-directed song. Since lullabies contain more pronounced versions of the salient characteristics of infant-directed speech, they might have an even more powerful impact on the infant.

Lullabies are found in the folk traditions of all countries and in the art music of all periods (New Grove Dictionary of Music and Musicians, 1980) but there has been no psychological research on their form and function. In collaboration with a musicologist (Anna Unyk), we are embarking on a series of descriptive and experimental studies of lullabies across a number of cultures. We will attempt to describe aspects of lullaby form that distinguish it from adult-directed song, identifying universal as well as culture-specific features. Although efforts to document common musical features across cultures have been relatively unsuccessful (see Harwood, 1976), we anticipate greater success with infant-directed song, in line with the cross-cultural similarities in infant-directed speech (Fernald et al., 1989). Our descriptive research will be supplemented by experimental research aimed at establishing whether infants exhibit a preference for infant-directed over adult-directed music and whether such musical forms have differential effects on attention and affect. Documenting the form of infant-directed music across cultures and its impact on infants will enrich our knowledge of the social and emotional context of early life, shedding light, perhaps, on musical universals in parenting.

References

Abberton, E., & Fourcin, A. J. (1978). Intonation and speaker identification. *Language and Speech, 21*, 305-318.

Alberts, J. R. (1981). Ontogeny of olfaction: Reciprocal roles of sensation andbehavior in the development of perception. In R. N. Aslin, J. R. Alberts, & M. R. Petersen (Eds.), *Development of Perception. Vol. I: Audition, somatic perception, and the chemical senses* (pp. 321-357). NY: Academic.

Ashmead, D., Clifton, R. K., & Perris, E. (1987). Precision of auditory localization in human infants. *Developmental Psychology, 23*, 641-647.

Attneave, F., & Olson, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology, 84*, 147-166.

Bahrick, L. E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development, 6*, 429-451.

Balch, W. R. (1981). The role of symmetry in the good continuation ratings of two-part tonal melodies. *Perception & Psychophysics, 29*, 47-55.

Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind, and brain* (pp. 321-351). NY: Plenum.

Banks, M. S., & Bennett, P. J. (1988). Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. *Journal of the Optical Society of America, 5*, 2059-2079.

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

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. (1989, April). *Phonological influence on infants' perception of two nonnative speech contrasts.* Paper presented at the Society for Research in Child Development, Kansas City, MO.

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

Bettes, B. A. (1988). Maternal depression and motherese: Temporal and intonational features. *Child Development, 59*, 1089-1096.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16*, 485-518.

Bharucha, J. J., & Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception & Psychophysics, 40*, 137-141.

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology, 6*, 145-238.

Bornstein, M. H., & Krinsky, S. J. (1985). Perception of symmetry in infancy: The salience of vertical symmetry and the perception of pattern wholes. *Journal of Experimental Child Psychology, 39*, 1-19.

Bregman, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 99-118). Hillsdale, NJ: Erlbaum.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89*, 244-249.

Bregman, A. S., & Dannenbring, G. L. (1973). The effect of continuity on auditory stream segregation. *Perception & Psychophysics, 13*, 308-312.

Butterworth, G., & Castillo, M. (1976). Coordination of auditory and visual space in newborn human infants. *Perception, 5*, 155-160.

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

Cernoch, J. M., & Porter, R. H. (1985). Recognition of maternal axillary odors by infants. *Child Development, 56*, 1593-1598.

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

Chang, H. W., & Trehub, S. E. (1977b). Infants' perception of temporal grouping in auditory patterns. *Child Development, 48*, 1666-1670.

Clarkson, M. G., Clifton, R. K., & Morrongiello, B. A. (1985). The effects of sound duration on newborns' head orientation. *Journal of Experimental Child Psychology, 39*, 20-36.

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

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

Corballis, M. C., & Beale, I. L. (1976). *The psychology of left and right.* Hillsdale, NJ: Erlbaum.

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

Cuddy, L. L. & Badertsher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics, 41*, 609-620.

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

de Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language, 11*, 1-15.

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

DeCasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: Preference, discrimination, and reinforcing value. *Developmental Psychology, 17*, 481-491.

DeCasper, A. J., & Sigafoos, A. D. (1983). Intrauterine heartbeat: A potent reinforcer for newborns. *Infant Behavior and Development, 6*, 19-25.

DeCasper, A. J., & Spence, M. (1986). Newborns prefer a familiar story over an unfamiliar one. *Infant Behavior and Development, 9*, 133-150.

Delack, J. B., & Fowlow, P. J. (1978). The ontogenesis of differential vocalization: Development of prosodic contrastivity during the first year of life. In N. Waterson & C. Snow (Eds.), *The development of communication* (pp. 93-110). Chichester: Wiley.

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

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature, 266*, 718-719.

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

Dowling, W. J. (1972). Recognition of melodic transformations. *Perception & Psychophysics, 12*, 417-421.

Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics, 14*, 37-40.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*, 429-354.

Dowling, W. J., & Fujitani, D. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America, 49*, 524-531.

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

Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research, 20*, 766-780.

Eimas, P. D., & Miller, J. L. (1980). Contextual effects in infant speech perception. *Science, 209*, 1140-1141.

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

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. (in press). Meaningful melodies in mother's speech to infants. In H. Papousek, V. Jurgens, & M. Papousek (Eds.), *Origins and development of nonverbal vocal communication: Evolutionary, comparative, and methodological aspects*. 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., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*, 104-113.

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

Field, J., Muir, D., Pilon, C., Sinclair, M., & Dodwell, P. (1980). Infants' orientation to lateral sounds from birth to three months. *Child Development, 51*, 295-298.

Fifer, W. P., & Moon, C. (1988). Auditory experience in the fetus. In W. Smotherman & S. Robinson (Eds.), *Behavior of the fetus* (pp. 175-188). West Caldwell, NJ: Telford.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). NY: Academic.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist, 58*, 34-42.

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

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.

Grieser. D. L., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology, 25*, 577-588.

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

Hebrank, J., & Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane. *Journal of the Acoustical Society of America, 56*, 1829-1834.

Hoffman, W. C., & Dodwell, P. C. (1985). Geometric psychology generates the visual gestalt. *Canadian Journal of Psychology, 39*, 491-528.

Huggins, A. W. F. (1972). On the perception of temporal phenomena in speech. *Journal of the Acoustical Society of America, 51*, 1279-1290.

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

Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? *Cognitive Psychology, 8*, 1-31.

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

Jones, M. R. (1982). Music as a stimulus for psychological motion: Part II. An expectancy model. *Psychomusicology, 2*, 1-13.

Jusczyk, P. W. (1989, April). *Perception of cues to clausal units in native and non-native languages.* Paper presented at the Society for Research in Child Development. Kansas City, MO.

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

Kessen, W., Levine, J., Wendrich, K. A. (1979). The imitation of pitch in infants. *Infant Behavior and Development, 2*, 93-99.

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

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

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

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

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

Kuhl, P. K. (1985). Methods in the study of infant speech perception. In G. Gottlieb & N. A. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 223-251). Norwood, NJ: Ablex.

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

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

Macfarlane, J. A. (1975). Olfaction in the development of social preferences in the human neonate. In M. A. Hofer (Ed.), *Parent-infant interaction.* Amsterdam: Elsevier.

Makin, J. W. & Porter, R. H. (1989). Attractiveness of lactating females' breast odors to neonates. *Child Development, 60*, 803-810.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79*, 487-509.

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

McAdams, S., & Bregman, A. S. (1979). Hearing musical streams. *Computer Music Journal, 3*, 26-43.

McNally, K. A., & Handel, S. (1977). Effect of element composition on streaming and the ordering of repeating sequences. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 451-460.

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

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

Mendelson, M. J., & Ferland, M. B. (1982). Auditory-visual transfer in four-month old infants. *Child Development, 53*, 1022-1027.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89-115.

Miller, J. L. (1977). Properties of feature detectors for VOT: The voiceless channel of analysis. *Journal of the Acoustical Society of America, 62*, 641-648.

Moore, B. C. J., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *Journal of the Acoustical Society of America, 80*, 479-483.

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.

Morrongiello, B. A. (1987). Infants' localization of sounds in the median sagittal plane: Effects of signal frequency. *Journal of the Acoustical Society of America, 82*, 900-905.

Morrongiello, B. A. (1988a). Infants' localization of sounds along the horizontal axis: Estimates of minimum audible angle. *Developmental Psychology, 24*, 8-13.

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

Morrongiello, B. A., & Clifton, R. K. (1984). Effects of sound frequency on behavioral and cardiac orienting in newborn and five-month old infants. *Journal of Experimental Child Psychology, 38*, 429-446.

Morrongiello, B. A., & Rocca, P. T. (1987). Infants' localization of sounds in the median vertical plane: Estimates of minimum audible angle. *Journal of Experimental Child Psychology, 43*, 181-193.

Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology, 44*, 413-426.

Muir, D. W. (1985). The development of infants' auditory spatial sensitivity. In S. E. Trehub & B. A. Schneider (Eds.), *Auditory development in infancy* (pp. 51-83). NY: Plenum.

Muir, D. W., & Clifton, R. (1985). Infants' orientation to the location of sound sources. In G. Gottlieb & N. A. Krasnegor (Eds.), *Measurement of audition and vision during the first year of postnatal life: A methodological overview* (pp. 171-194). Norwood, NJ: Ablex.

Muir, D. W., Clifton, R. K., & Clarkson, M. G. (1989). The development of a human auditory localization response: A U-shaped function. *Canadian Journal of Psychology, 43*, 199-216.

Muir, D. W. & Field, J. (1979). Newborn infants orient to sounds. *Child Development, 50*, 431-436.

New Grove Dictionary of Music and Musicians. (1980). (pp. 313-314). London: MacMillan.

Papousek, M, Papousek, 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: Ablex.

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

Papousek, M., & Papousek, H. (in press). Melodic units in maternal speech in tonal and nontonal languages: Evidence of a universal parental support for linguistic communication. *Infant Behavior and Development.*

Perris, E. E., & Clifton, R. K. (1988). Reaching in the dark as a measure of auditory localization in 7-month old infants. *Infant Behavior and Development, 11*, 477-495.

Querleu, D., & Renard, X. (1981). Les perceptions auditives du foetus humain. *Médecine et Hygiène, 39*, 2102-2110.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crepin, G. (1988). Fetal hearing. *European Journal of Obstetrics & Gynecology & Reproductive Biology, 29*, 191-212.

Roberts, L. A., & Shaw, M. L. (1984). Perceived structure of musical triads. *Music Perception, 2*, 95-124.

Rosch, E., & Mervis, C. B. (1975). Family resemblance studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Schenker, H. (1906/1954). *Harmony* (O. Jones, Ed., & E. M. Borgese, Trans.). Cambridge: MIT. (Original published 1906).

Simonton, D. K. (1984). Melodic structure and note transition probabilities: A content analysis of 15,618 classical themes. *Psychology of Music, 12*, 3-6.

Singh, P. G. (1987). Perceptual organization of complex tone sequences: A tradeoff between pitch and timbre? *Journal of the Acoustical Society of America, 2*, 886-899.

Spelke, E. S., Born, W. S., & Chu, F. (1983). Perception of moving, sounding objects by four-month old infants. *Perception, 12*, 719-732.

Spence, M. J., & DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal-voice sounds influences neonatal perception of maternal voice samples. *Infant Behavior and Development, 10*, 133-142.

Stack, D. M., Muir, D. W., Sherriff, F., & Roman, J. (1989). Development of infant reaching in the dark to luminous objects and 'invisible sounds'. *Perception, 18*, 69-82.

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.

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

Studdert-Kennedy, M. (1975). From continuous signal to discrete message: Syllable to phoneme. In J. F. Kavanaugh & J. E. Cutting (Eds.), *The role of speech in language* (pp. 113-125). Cambridge, MA: MIT.

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

Thorpe, L. A. (1985). *Auditory-temporal organization: Developmental perspectives.* Unpublished doctoral dissertation, University of Toronto.

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.

Tougas, Y., & Bregman, A. S (1985). The crossing of auditory streams. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 788-798.

Trehub, S. E. (1979). Reflections on the development of speech perception. *Canadian Journal of Psychology, 33*, 369-381.

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

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

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

Trehub, S. E. (1990). The perception of musical patterns by human infants: The provision of similar patterns by their parents. In M. Berkley & W. C. Stebbins (Eds.), *Comparative perception, Vol. 1: Basic mechanisms*. NY: 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., Cohen, A. J., Thorpe, L. A., & Morrongiello, B. A. (1986). Development of the perception of musical relations: Semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 295- 301.

Trehub, S. E., Schneider, B. A., & Endman, M. (1980). Developmental changes in infants' sensitivity to octave-band noises. *Journal of Experimental Child Psychology, 29*, 282-293.

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. (in press). Infants' perception of *good* and *bad* melodies. *Psychomusicology.*

Van Lanker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: patterns and parameters. Part I: Recognition of backward voices. *Journal of Phonetics, 13*, 19-38.

Watkins, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception & Psychophysics, 37*, 275-285.

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

Wishart, J. G., Bower, T. G. R., & Dunkfield, J. (1978). Reaching in the dark. *Perception, 7*, 507-512.

Wolff, P. H. (1987). *The development of behavioral states and the expression of emotions in early infancy.* Chicago: University of Chicago Press.

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

Zelazo, P. R., Brody, L. R., & Chaika, H. (1984). Neonatal habituation and dishabituation of head turning to rattle sounds. *Infant Behavior and Development, 7*, 311-321.

Acknowledgements

The preparation of this chapter was assisted by grants to the senior author from the Natural Sciences and Engineering Research Council of Canada and the University of Toronto. We thank A. S. Bregman for his constructive comments on an earlier draft of the chapter.