INFANT BEHAVIOR AND DEVELOPMENT 20 (3), 1997, pp. 383–396 ISSN 0163-6383

The Acoustic Basis of Preferences for Infant-Directed Singing

LAUREL J. TRAINOR, ELISSA D. CLARK, ANITA HUNTLEY, AND BETH A. ADAMS

McMaster University

Acoustic differences between infant-directed and noninfant-directed singing were examined in 6 playsong and 4 lullaby pairs of recordings from Trainor (1996). Each mother had been recorded singing a song of her choice to her infant and singing the same song in the absence of her infant. For both playsongs and lullabies the tempo was slower, there was relatively more energy at lower frequencies, inter-phrase pauses were lengthened, and the pitch and jitter factor were higher in the infant-directed over infant-absent versions. Pitch variability was higher and the rhythm exaggerated (as measured by the relative duration of stressed to unstressed syllables) in the infant-directed versions of playsongs but not lullabies. Many of these acoustic modifications likely attract infants' attention, and playsongs and lullabies likely communicate different emotional messages.

music song emotion acoustic mother–infant interaction rhythm pitch

Although the practice of singing to infants appears to be cross-culturally universal (Trehub, Unyk, & Trainor, 1993a, 1993b; Trehub, Trainor, & Unyk, 1993; Unyk, Trehub, Trainor, & Schellenberg, 1992), psychologists have only recently begun to study the nature and function of this activity. Previous research has indicated that there are special songs for infants that are perceptually distinct from adult-directed songs (Trehub et al., 1993a). The present paper concerns the acoustic characteristics of the special way or style of singing adopted by caregivers when interacting with infants. Many studies show that caregivers *talk* to preverbal infants in a special way (e.g., Fernald, 1991; Papousek, 1992) and that infants prefer to listen to this infant-directed speech over adult-directed speech (e.g., Cooper & Aslin, 1990; Fernald, 1993; Fernald & Kuhl, 1987). Across different languages and cultures, infant-directed speech tends to be higher in pitch, more rhythmic, and, most important for mediating infant preferences, contains slower, more exaggerated pitch contours than adult-directed speech. In a sense, then, infant-directed speech could be called musical speech, because it differs from adult-directed speech in its prosodic or musical characteristics. Furthermore, these studies suggest that communication between caregivers and preverbal infants takes place through musical features (Fernald, 1989; Papousek, Bornstein, Nuzzo, Papousek, & Symmes, 1990; Werker & McLeod, 1989).

Trehub et al. (1993b) first demonstrated that infant-directed singing differs from singing in the absence of an infant (infant-absent singing). They recorded English-speaking and Hindi-speaking mothers singing a song of their choice to their baby and singing the same song in the absence of their baby. Although there were some differences across the two cultures (e.g., Hindi mothers tended to sing soothing songs whereas English-speaking mothers tended to sing rousing playsongs), adult raters from both cultures were able to distinguish the infant-directed from the infant-absent versions. These results were replicated by Trainor (1996, Experiment 1) for English-speaking mothers. In a further study (Trehub et al., in press), mothers' singing to their infants was contrasted with their attempts at singing as if to their infant when their infant was not actually present. A variety of adult raters, from university students to mothers to graduate students specializing in early childhood education, were able to distinguish the infant-directed from the simulated samples. Furthermore, fathers, who generally have less caretaking experience than mothers, also make similar modifications in their singing to infants (Trehub et al., in press) as do children (Trehub & Schellenberg, in press), suggesting that extensive experience with infants is not required for

Direct all correspondence to: Laurel J. Trainor, Department of Psychology, McMaster University, Hamilton, ON, Canada L8S 4K1; e-mail: ljt@mcmaster.ca.

such modification to occur. The fact that the infant needs to be present for the full range of infant-directed features to appear suggests that these modifications are intuitive and that mothers are to some extent unaware of them.

Trainor (1996) suggested that infant-directed singing might serve a number of purposes or functions. One function might be simply to attract infants' attention to an important person in the environment, the caregiver. Trainor (1996) recorded 15 mothers singing a song of their choice to their infant (infant-directed version) and in the absence of their infant (infant-absent version). She subsequently presented the songs from 6 mothers to infants in a looking-time preference procedure (Experiment 2). For 5 of the 6 pairs, infants looked significantly longer to produce the infant-directed version than they did to produce the infant-absent version, indicating that the infant-directed versions attract infants' attention.

A second function of infant-directed singing as outlined by Trainor (1996) concerns emotional communication and the regulation of the infant's state. Music is closely associated with emotion (e.g., Bever, 1988; Cooke, 1959; Langer, 1957; Meyer, 1956). The contour of tension build-up and release that occurs as a musical piece unfolds has been described as a language of emotions (Cooke, 1959; Langer, 1957). Music can both communicate information about emotion and evoke direct emotional responses, as evidenced by reactions such as shivers down the spine, laughter, tears, and lump in the throat (Sloboda, 1991). Not surprisingly, across many societies music is associated with magical powers and is used in medicine and healing. In Western society, this field is called music therapy (Davis, Gfeller, & Thaut, 1992). Further, children at least as young as 3 or 4 years old are sensitive to emotional meaning in music (Cunningham & Sterling, 1988; Gentile, Pick, Flom, & Campos, 1994; Trainor & Trehub, 1992). Early interactions between caregivers and their infants are largely concerned with providing comfort and easing unhappiness. Thus, caregivers are intimately concerned with helping regulate their infant's state. Music, whether in the form of song or musical speech, would appear to be the verbal form of communication most suited to this purpose. Certainly, infants' preference for musical, or infant-directed, speech over adult-directed speech suggest this to be the case, and infants as young as 5 months old are sensitive to emotional meanings conveyed by speech prosody (Caron, Caron, & MacLean, 1988; Walker-Andrews & Grolnick, 1985).

Trainor (1996, Experiment 1) obtained a number of adult ratings of the 15 infant present/infant absent pairs of mothers' singing. A first group of adults was 93% correct at identifying the infant-directed over the infant-absent versions. A second group rated the infant-directed versions as being in a more loving tone of voice 83% of the time, suggesting that for adult raters the infant-directed versions were more emotionally charged. Infants were also sensitive to this variable as the degree of infant preference, as measured by looking time (Experiment 2), was significantly correlated with the adult ratings of loving tone of voice (Experiment 1). In fact, the loving or smiling tone of voice has surfaced as an important variable across several studies of infant-directed singing (Trainor, 1996; Trehub et al., 1993b, in press). A third group of adult raters from Trainor (1996, Experiment 1) consistently classified six of the infant-directed versions as playsongs and four as lullabies. The infant-directed versions of the playsongs were rated as more rhythmic than their infant-absent pair, whereas the opposite was true for the lullabies. Thus, caregivers appear to sing differently in different caretaking situations, with one type of singing to calm infants and induce sleep and another type of singing to arouse and engage infants in play.

The third function of infant-directed singing suggested by Trainor (1996) concerns teaching infants about auditory pattern structure, that is, about phrase structure, rhythm, and grouping. Infants are sensitive to phrase structure in music (e.g., Jusczyk & Krumhansl, 1993) and infant-directed speech (e.g., Jusczyk et al., 1992). They are also able to discriminate different rhythmic structures (Allen, Walker, Symond, & Marcell, 1977; Demany, McKenzie, & Vurpillot, 1977; Mendelson, 1986; Morrongiello, 1984; Trehub & Thorpe, 1989). It has been proposed that one of the functions of the "musicalization" of infant-directed speech is to exaggerate structural characteristics, thereby helping infants learn language (e.g., Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989). For instance, Fernald and Mazzie (1991) found that the most important words of a phrase were uttered at the highest pitch level. Bernstein Ratner (1986) found that mothers exaggerated cues to clause boundaries in their speech to infants: clause-final vowels were lengthened relatively more in infant-directed than adult-directed speech. Further, infants may initially gain access to word boundary segmentation by learning the stress or rhythmic patterns—again, a musical type of feature—of the language they are learning (e.g., Echols, 1996; Jusczyk, Cutler, & Redanz, 1993; Morgan, 1994). Following this logic, early experience with music itself might help develop the general auditory pattern processing skills necessary for decoding both speech and music.

The present paper is concerned with identifying the acoustic differences between infant-directed singing and singing in the absence of an infant. The overriding interest was in the function of infant-directed singing. Why do caregivers around the world sing to their infants? Why do North American mothers, many of whom would otherwise never sing, sing to their infants? In particular, the six playsong and four lullaby pairs consistently identified as such in Trainor (1996) formed the material for the acoustic analyses reported here. The choice of the particular acoustic measures was driven by the three hypotheses for the function of infant-directed singing outlined above.

If the purpose of infant-directed singing is to attract infants' attention, a number of acoustic differences might be expected between infant-directed and infant-absent versions. Infants prefer the higher pitch and exaggerated prosody of infant-directed speech, particularly as manifested in increased pitch contours and elongated vowels (Fernald & Kuhl, 1987). Indeed, previous research has found that infant-directed singing also tends to be higher in pitch than infant-absent singing (Trehub et al., in press). While the tune of the song obviously constrains the modification of pitch contour, increased variation in other dimensions, such as dynamic range (i.e., intensity difference between loudest and softest sounds), accentuation (e.g., relative length and intensity of stressed to unstressed syllables), tempo (i.e., rate), and the relative length of vowels within syllables might characterize the infant-directed singing versions relative to the infant-absent versions.

With respect to the second possible function of infant-directed singing, the emotional content of the singing samples was examined with a number of acoustic measures. As discussed by Scherer (1986) and Frick (1985), listeners easily distinguish vocal emotions in speech, but the acoustic correlates of specific emotions are very difficult to describe. The vast majority of studies of emotion in voice have used speech rather than sung materials. According to Scherer's (1986) component process model, emotions arise through the processing and interpretation of information. The results of continual stimulus evaluation checks are expressed physiologically through changes in the autonomic and somatic nervous systems. Such changes affect the musculature involved in vocal expression and lead to the acoustic differences perceived in the expression of different emotions. For example, when pleasantness is experienced, the faucal and pharyngeal muscles expand to produce a "wide voice," and the vocal tract is shortened due to the retraction of the corners of the mouth. The former has the effect of concentrating relatively more energy at lower frequencies, and the latter has the effect of raising the resonant (i.e., formant) frequencies of the vocal tract.

According to Murray and Arnott (1983), people perceive emotion in the voice primarily through *voice quality* (Lieberman and Michaels, 1962), *utterance timing* (Knower, 1941), and *utterance pitch contour*. In music, the pitch contour is largely constrained by the musical structure. However, if mothers are conveying emotional meaning to their infants, then differences in voice quality and utterance timing would be expected between infant-directed and infant-absent versions.

Of the primary emotions studied in adulthood (anger, joy/happiness/humor, sadness, fear/anxiety, and disgust/hatred/contempt/scorn; see Murray & Arnott, 1983) joy/happiness probably best describes the emotions expressed in singing playsongs to infants. To adequately describe the emotions associated with lullabies, however, it is necessary to examine the secondary emotions. Of these (grief/sorrow, affection/tenderness, sarcasm/irony, and surprise/astonishment), affection/tenderness probably comes closest to those experienced when singing a lullaby. While the emotional content of playsongs and lullabies may be somewhat different from the adult emotions of joy/happiness and affection/tenderness, acoustic analyses of these emotions can serve as guides to the choice of acoustic parameters that might differ between infant-present and infant-absent versions of playsongs and lullabies.

The emotion of joy is associated with increased pitch or fundamental frequency (Scherer, 1986) and more variability in fundamental frequency (Scherer, 1986), so these characteristics were measured in our sample of lullabies and playsongs. Joy is also associated with increased intensity range (Scherer, 1986). This was measured by examining the relative intensity differences between stressed and unstressed syllables. Compared to neutral speech, the relative proportion of energy in the higher frequency range is reported to increase in joy but decrease in happiness (Scherer, 1986). The tempo or rate of speech is reported to be faster in joy, but slower in happiness (Scherer, 1986; see also Murray & Arnott, 1983). Finally, jitter (variation in the fundamental frequency at the smallest time period) and shimmer (variation in intensity at the smallest time period) are expected to increase with the amount of emotional involvement (e.g., Bachorowski & Owren, 1995).

Very little is known about the acoustic correlates of affection/tenderness. With this emotion, fundamental frequency has been reported to be both higher (Fonagy & Magdics, 1963) and lower (Davitz, 1964) than in neutral speech (see Murray & Arnott, 1983). Affection/tenderness has also been characterized as being slower in tempo (Fonagy & Magdics, 1963; Davitz, 1964) and with a more regular rhythm (Davitz, 1964). Thus, we compared the infant-directed and infant-absent versions on tempo and variation in tempo. From the characterization of affection/tenderness as having a "full" or "resonant" timbre (Fonagy & Magdics, 1963; Davitz, 1964), we expected the relative proportion of energy in the lower frequency range to be higher for the infant-directed over infant-absent lullaby samples.

Finally, if one of the functions of music in infancy is to give infants experience with the structure of temporally extended sound patterns, one might expect mothers singing to their infants to intuitively exaggerate the structure of the songs when singing to their infant. Indeed, a couple of studies have found that mothers sing more slowly to their infants than otherwise, suggesting that such a didactic function might be involved (Trehub et al., 1993b, in press). If mothers were exaggerating the phrase structure, we hypothesized that they might increase the length of phrase-final syllables and pauses between phrases. Possible exaggeration of the rhythmic structure was measured by comparing the relative duration and intensity of stressed to unstressed syllables across the infant-directed and infant-absent versions.

As can be observed from the above discussion, the acoustic parameters associated with the three possible functions of infant-directed speech, attracting infants' attention, state regulation and communicating information about emotion, and teaching infants about auditory pattern structure, overlap to some extent. The parameters studied were split into two main categories: those related to voice quality and those related to the clarity of musical structure.

METHOD

Participants

The majority of the results reported here are measures of acoustic parameters, performed by a technician with considerable experience in this domain. A second technician duplicated a subset of the syllable duration measurements to ensure that a consistent criterion was being applied. Pitch variability was studied through adult ratings. The 10 adults in question were either graduate students in psychology or had some knowledge of acoustics.

Apparatus

The acoustic analyses were performed with two commercial software systems. In one case, the tape recorded singing samples from Trainor (1996) were played through a Tucker Davis amplifier (MA2), digitized through a Tucker Davis analog interface (DD1) (sampling rate 22,000 Hz) connected to a Comptech 486 computer with a Tucker Davis AP2 processor card and the Computerized Speech Research Environment (CSRE) software package. In the other case, the recordings were digitized through an Audiomedia II card in a Macintosh IIci computer running the Signalyze 3.09 speech analysis software program.

Stimuli

Six of the infant-directed songs recorded by Trainor (1996) were rated as playsongs by 90% or more of adults, and four were rated as lullabies by 80% or more of adults. These 10 infant-directed/infant-absent pairs were the material for the acoustic analyses of the present report (see Table 1). In each case, mothers were recorded in a quiet laboratory room sing-

TABLE 1
Phrases, Stressed Syllables, And Beats Of The Songs Analyzed

Playsongs

Mother 1
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU ////
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU
I //// LOVE you / in the / morn- / ing and / IN the / after- / noon
I // LOVE you / in the / eve- / ning / UNDER- / neath the / moon //
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU ////

Mother 2
EENsy /weensy / spi- / der went / UP the / water / spout //
DOWN / came the / rain / and / WASHED the / spider / out //
OUT / come the / sun / and / DRIED up / all the / rain / so /
EENsy / weensy / spi- / der climbed / UP the / spout a- / gain //

Mother 3
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU ////
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU
I //// LOVE you / in the / morn- / ing and / IN the / after- / noon
I // LOVE you / in the / eve- / ning / UNDER- / neath the / moon //
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU ////

Mother 4
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU
I //// LOVE you / in the / morn- / ing and / IN the / after- / noon
I // LOVE you / in the / eve- / ning / UNDER- / neath the / moon //
SKI-da-mar / ink-i / dink-i / dink / SKI-da-mar / ink-i / do //
I // love // YOU ////

Mother 5
ROW / row / row your / boat /
GENtly / down the / stream //
MERRily / merrily / merrily / merrily /
LIFE is / but a / dream //

Mother 6
BAA / baa / black / sheep / HAVE you / any / wool //
YES / sir / yes / sir / THREE / bags / full //
ONE / for my / mas- / ter and / ONE / for my / dame //
ONE / for the / little / boy who / LIVES / down the / lane //
BAA / baa / black / sheep / HAVE you / any / wool //
YES / sir / yes / sir / THREE / bags / full //

Lullabies

Mother 7
You / are / my / SUN // shine ///
My / on- / ly / SUN // shine ///
You / make / me / HA- // ppy ///
When / skies / are / GREY /////
You'll / ne- / ver / know // DEAR ///
How / much / I / LOVE // you ////
Please / don't / TAKE ///
My / sun // shine / a- / WAY ////////

TABLE 1
Continued

Mother 8
ROCK-a- // bye / ba- // by / ON the // tree / top ///
WHEN the // wind / blows // the / CRA- / dle / will / rock ///
WHEN the // bow / breaks // the / CRA- / dle / will / fall
And /// DOWN / will / come / Ma- // ry / CRA- / dle / and / all ///

Mother 9
PUFF the / magic / dra- / gon /
LIVED / by the / sea
And // FROlicked / in the / autumn / mist
In a / LAND called / hona- / lee //
PUFF the / magic / dra- / gon /
LIVED / by the / sea
And // FROlicked / in the / autumn / mist
In a / LAND called / hona- / lee //

Mother 10
I / love / you // YOU / love / me //
WE'RE / a / ha- / ppy / FA- / mi- / ly /
With a / GREAT / big / hug / and a / KISS from / me to / you //
WON'T / you / say / you / LOVE / me / too //

Note. Each line represents one phrase. Secondary stresses are underlined; primary stressed syllables are capitalized and underlined. Beats are separated by slashes. When one syllable is more than one beat in duration, the number of slashes after that syllable indicates the number of beats.

ing the same song to their infant and in the absence of their infant. Mothers who sang first to their infant were simply told we wanted to collect an additional recording (without their infant). Mothers who sang alone first were simply asked to sing a song they normally sang to their infant and subsequently to sing the same song to their infant. Most mothers held their infants, but an infant seat was available.

Procedure

The variables analyzed fall into two main categories, those related to *voice quality* and those related to the *musical structure*. In each case, measurements were taken on the identical phrases or portions thereof sung by the same mother in presence and absence of her infant. In cases where a reliable measurement could not be taken in one version (e.g., a portion of the signal was masked by baby vocalization), that measurement was eliminated from both versions. A number of measures were performed on the stressed and unstressed syllables. The stressed syllables were determined by a professional musician who was familiar with the songs but had not heard the recordings (see Table 1). In some cases only the primary stressed syllables were used, and in other cases both the primary and secondary stressed syllables were used, as detailed below.

Measures Related to Voice Quality

1. *Mean pitch.* The pitch of the centre of the steady state portion of the vowel in all primary stressed syllables was measured using the Cepstrum Function in the CSRE system. Calculation windows were 23.3 ms and were overlapped 60%. The data from one mother in the lullaby category (Mother 10) was eliminated as the infant-directed version was almost whispered, and it was impossible to measure the pitch reliably. The pitch of a further 4 of the 94 vowels (2 from lullabies and 2 from playsongs) could not be measured due to infant vocalization. The pitch values obtained for the rest of the singing samples were checked against pitch measurements obtained with the Comb Filtering Function. In this way, one octave error was corrected (Mother 5), and two vowels were eliminated (from lullabies) because the two methods yielded pitches differing by more than 30 Hz.

2. *Jitter or frequency perturbation.* The jitter factor measures variation in the fundamental frequency at the smallest time period. Specifically, it is the mean difference between the frequencies of adjacent periods (phonation cycles) divided by the mean frequency, multiplied by 100 (see Baken, 1987, p. 175). Jitter was measured for the same vowels of the primary stressed syllables as in the fundamental frequency analysis above, using the built-in function in the CSRE system.

3. *Shimmer or amplitude perturbation.* Shimmer is similar to jitter, except that the dB difference between peak amplitudes of adjacent cycles is the measure of interest (see Baken, 1987, p. 116). Shimmer was measured for the same vowels of the primary stressed syllables as in the fundamental frequency and jitter analyses above, using the built-in function in the CSRE system.

4. *Ratings of pitch variability.* In addition to the short-term pitch variability measured by jitter, we were interested in longer-term pitch variability. Rather than devise a new

mathematical formulation to capture this idea, which would have then needed extensive validation, we chose to visually present graphs of pitch extractions of identical excerpts of infant-directed and infant absent pairs to adult raters, who were asked to choose the version in which the pitch looked most variable. (Adult raters were not informed that the pairs consisted of one infant-directed and one infant-absent version). Pitch extractions were performed with CSRE's cepstrum function. Pitch was calculated with 23.3 ms windows that overlapped by 60%. It was only possible to obtain clear pitch extractions across phrases for 1 of the 4 lullabies ("Rock-a-bye baby"), as mothers sang too softly, or whispered parts of the song in the other 3 cases. The 3 playsongs showing the clearest pitch extractions were also included ("Baa baa black sheep" and 2 versions of "Skidamarink"). In all, raters were presented with 16 pairs of pitch extractions for each "Skidamarink" (corresponding to phrases), 12 pairs for "Baa baa black sheep" (corresponding to half-phrases), and 6 pairs for "Rock-a-bye baby" (corresponding to 2 phrases and 4 half-phrases). Initially, participants were given one example with detailed instructions.

5. *Relative intensity of low, mid, and high frequencies.* Formant frequencies were measured initially. However, the use of different algorithms led not only to different absolute values of the formant frequencies, but the relative ranking of the infant-present and infant-absent versions often reversed. This measurement difficulty probably stems from several factors: the pitch was relatively high (resulting in widely space harmonics), the samples were sung not spoken, and mothers tended to use a large intensity range especially when singing to their infants.

Instead of formant frequency, we used a more crude measure of the spread of intensity over frequency. Using the vowels of the primary stressed syllables, we compared the intensity in three frequency bands (using the Signalyze program). The signal was filtered using a Butterworth filter (order 10) in three ways: low pass with 600 Hz cutoff (capturing the fundamental frequency), band pass between 600 and 2000 Hz, and band pass between 2000 and 4000 Hz. For each resultant waveform, a RMS envelope was calculated with a window width of 20 ms, and the average RMS for 60 ms selected from the centre of the vowel was used. The dB difference between the lowest frequency band and each of the higher two bands were calculated using the formula $20\log_{10}(RMS_{band}1/RMS_{band2})$. Three of the 102 vowels (2 from playsongs and 1 from a lullaby) could not be included due to infant vocalization.

Measures Related to the Clarity of Musical Structure

In order to compare the clarity of the musical structure in the infant-present and infant-absent renditions, we examined acoustic features related to rhythm. All durations were measured in Signalyze. Signals were initially low pass filtered with a cutoff of 5000 Hz using an order 10 Butterworth filter. Two displays were examined simultaneously: the amplitude over time waveform and a spectrogram calculated using a very wide band (suitable for high pitched voices) with pre-emphasis and smoothing. Decisions concerning where one syllable ended and the next began were based primarily on the waveform, but confirmed by examining the spectrogram. As the exact determination of such boundaries is subject to interpretation, a second technician duplicated some of the duration measurements for one playsong and one lullaby pair after having been briefed as to the precise criteria being used. For the playsong, the second technician measured the duration of 35 stressed syllables for each of the infant-present and infant-absent pairs. The correlation between the duration measurements of the two technicians were .999 and .997 for the infant present and infant absent durations, respectively. Given this high level of agreement, only six stressed syllables were measured in each case for the lullaby. Correlations in this case were .999 and .996, respectively.

1. *Timing deviation at the ends of phrases: Length of inter-phrase pauses and phrase-final syllables relative to phrase length.* Longer silent pauses between phrases can make the phrase structure of the music perceptually clearer. As the infant-directed and infant-absent versions were at different tempos, the duration of each inter-phrase pause was measured as a percentage of the duration of the preceding phrase (excluding the final syllable and pause between phrases). To see whether increased silent pauses disrupted the rhythm or whether mothers compensated for any increased pause duration with decreased final syllable duration, the length of the final syllable plus inter-phrase pause was measured as a percentage of the duration of the preceding phrase. Final phrases of each song could not be used as there is by definition no following inter-phrase pause. An additional 6 phrases were also eliminated from the playsong pairs, 2 because infant vocalizations precluded precise duration measurement and 4 because the mother shortened the final syllable in the infant-directed version in order to insert extra nonsense syllables. Three phrases were eliminated from the lullaby pairs due to infant vocalizations.

2. *Accent structure: Relative intensity of stressed and unstressed syllables.* Intensity is one of the acoustic correlates of stress or accent. This measure gives an indication of the dynamic range (intensity variability) and is probably related to the perception of rhythmicity. The dB difference between each primary stressed syllable and a nearby unstressed syllable was calculated as in Number 5, above. Where possible, the syllable following the stressed syllable was used as the unstressed comparison. If the following syllable was also stressed, the next closest unstressed syllable was chosen, whether it was before or after the stressed syllable. In 15 cases in the playsongs the unstressed syllable was from the following phrase. In addition, for 3 of the 70 stressed syllables from the playsongs and 1 of the 32 from the lullabies it was not possible to take measurements due to infant vocalization.

3. *Accent structure: Relative duration of stressed and unstressed syllables.* Duration is one of the acoustic correlates of stress or accent. The ratio of the duration of stressed and unstressed syllables for primary and secondary stresses were calculated. This measure gives an indication of the relative degree of accent put on the stressed syllables through lengthening. Again, if the following syllable was also stressed, the next closest unstressed syllable was chosen, whether it was before or after the stressed syllable. In 22 cases in the playsongs and 5 cases in the lullabies, the unstressed syllable was from the following phrase. For 5 of the 102 primary stressed syllables (4 playsongs; 1 lullaby) and 9 of the 99 secondary stressed syllables (4 playsongs; 5 lullabies), it was not possible to obtain measurements due to infant vocalization on either the stressed or unstressed syllable.

4. *Accent structure: Stressed syllable duration relative to phrase length.* The ratio of the duration of each primary and secondary stressed syllable to the phrase in which it occurred was calculated. This measure gives another indication of the relative degree of stress present in the singing. For 7 of the 132 primary stressed syllables (3 in playsongs; 4 in lullabies) and 6 of the 99 secondary stressed syllables (3 in playsongs; 3 in lullabies), it was not possible to obtain measurements due to infant vocalization.

5. *Relative duration of the vowel in stressed syllables.* In infant-directed speech, vowels are often elongated (e.g., Furguson, 1964). Musical rhythm is more constraining than speech rhythm, but it is possible that the vowels of primary stressed syllables might be lengthened relative the length of the syllable in which they occur. Accordingly, for each primary and secondary stressed syllable used in Number 4, above, we measured the ratio of the duration of each vowel relative to the duration of the syllable.

6. Tempo. The main beats of each song were determined by a professional musician (see Table 1) and the onset-to-onset duration of each beat measured. Phrase-final beats were excluded. In addition, infant vocalizations precluded obtaining measurements from 12 of the 108 beats in the playsong pairs and 10 of the 121 beats in the lullaby pairs. In cases where a syllable was held over a beat boundary, the duration of the two (or more) beats was measured, and the value divided by the number of beats, yielding an average beat length for the two (or more) beats. The tempo is represented by the average beat duration.

7. *Tempo variability.* For each mother, the average beat duration was determined, and the variation in tempo was examined by calculating the absolute value of the difference between each beat and it's expected value (average beat duration).

RESULTS

The results are summarized in Table 2. For each mother, the same measurements were taken on the corresponding components in the infant-directed and infant-absent renditions of the same song. As different numbers of measurements were taken with each mother, depending on the length of the song, for each acoustic measure a standard score was calculated for each mother that represented the magnitude of difference between the infant-directed and infant-absent versions. Specifically, the standard score for each mother was the average difference between infant-directed and infant-absent paired scores divided by the standard deviation. Differences between infant-directed and infant-absent versions were examined with t tests. Under the null hypothesis, the expected value of the standard score is 0. Unpaired two-group t tests were conducted on the standard scores to examine differences between lullabies and playsongs.

Measures Related to Voice Quality

1. *Mean pitch.* For both lullabies and playsongs, the mean pitch was significantly higher in the infant-directed versions, $t(5) = 5.99, p < .001$ for playsongs and $t(2) = 3.12, p < .05$ for lullabies, and every mother showed this trend. Relative to the infant-absent versions, the pitch was raised more in the infant-directed playsongs than the infant-directed lullabies, $t(7) = 2.52, p < .02$ (mean of 2.92 semitones and 2.44 semitones for playsongs and lullabies, respectively). Overall, the playsongs were sung at a higher pitch level than the lullabies (mean in Hz = 253.6 for playsongs, 224.5 for lullabies).

TABLE 2
Acoustic differences between infant-directed and infant-absent singing.

Acoustic Parameter	Playsongs	Lullabies
Mean pitch	+	+
Jitter	+	+
Shimmer	+	
Pitch variability	+	
Relative intensity of low frequencies	+	+
Relative length of inter-phrase pauses	+	+
Relative length of phrase-final syllables + pauses	+	
Relative intensity of stressed to unstressed syllables		
Relative duration of stressed to unstressed syllables	+	
Stressed syllable duration relative to phrase		−
Relative vowel duration in stressed syllables		
Tempo (beat duration)	−	−
Tempo variability		

Note. Plus signs indicate that the acoustic parameter was significantly greater in the infant-directed versions, minus signs that it was greater in the infant-absent versions, and no entry that there was no difference between infant-directed and infant-absent versions.

2. *Jitter or frequency perturbation.* The jitter factor was higher in the infant-present than the infant-absent versions of both the playsongs, $t(5) = 5.32$, $p < .002$, and the lullabies, $t(2) = 3.90$, $p < .03$, and the difference between playsongs and lullabies was not significant. For playsongs, the mean jitter factor was 22.32 for infant present compared to 13.90 for infant absent versions. For lullabies, these figures were 24.51 and 13.96, respectively.

3. *Shimmer.* Lullabies and playsongs were similar in the amount of shimmer increase from infant-directed to infant-absent versions (means in dB of 3.60 and 2.14 for playsongs and 4.34 and 3.87 for lullabies). Together, shimmer was significantly greater in the infant-directed over the infant-absent versions, $t(8) = 2.74$, $p < .013$. This difference reached conventional levels of significance for the playsongs alone, $t(5) = 2.27$, $p < .04$, although not for lullabies, $p > .15$.

4. *Ratings of pitch variability.* The percentage of trials on which adults rated the infant-directed visual display of each phrase as more variable in pitch than the infant-absent visual display was calculated for each rater for each mother. For each of the three playsongs, these percentages were significantly above 50%, indicating that the infant-directed pitch contours were rated as more variable than their infant-absent pairs [$t(9) = 9.80$, $p < .0001$, $t(9) = 15.09$, $p < .0001$ for the two mothers who sang "Skidamarink", and $t(9) = 3.08$, $p < .013$ for "Baa baa black sheep"]. (Even with a Bonferroni correction, these two-tailed tests are all significant at the .05 level). An ANOVA revealed that there was a difference across mothers, $F(2,27) = 28.19$, $p < .0001$, and Scheffe *F*-tests revealed that the percentage of phrases in which the infant-directed version was rated as more variable than the infant-absent version did not differ between the two mothers who sang "Skidamarink" (mean = 80.0 and 85.0%), but this percentage was significantly lower for the mother who sang "Baa baa black sheep" (mean = 58.1%) than either of the "Skidamarink" renditions. By contrast, for the lullaby the percentage of trials on which the infant-directed version was rated as more variable did not differ significantly from 50%, $t(9) = -1.15$, $p > .28$ (mean = 45%). It is hard to generalize with too much confidence from one lullaby, but the results suggest that pitch contour is more variable in infant-directed than infant absent playsongs, while infant-direct and infant-absent lullabies do not differ in this respect.

5. *Relative intensity of low, mid, and high frequencies.* The dB difference between the low-frequency band (< 600 Hz) and the mid-frequency band (600–2000 Hz) for the infant-directed and infant-absent versions was greater for lullabies than for playsongs, $t(8) = 1.92$, $p < .05$. This difference was not significant across infant-directed and infant-absent versions of playsongs, $p > .5$, but approached conventional levels of significance for lullabies, $t(3) = 1.90$, $p < .08$, suggesting that there was relatively more energy at low than mid frequencies in infant-directed lullabies. It is possible that this difference resulted because mothers tend to hold their baby lower and look down more when singing lullabies than playsongs, resulting in a more closed jaw. Whatever the origin, however, the acoustic differences and their perceptual effects remain. The dB differences between the low-frequency band and the high-frequency band (2000–4000 Hz) did not differ between playsongs and lullabies. Together, there was relatively more energy at low frequencies in infant-directed versus infant-absent versions, $t(9) = 2.59$, $p < .02$, and separately, this factor approached significance for both playsongs, $t(5) = 1.87$, $p < .06$, and lullabies, $t(3) = 1.67$, $p < .10$.

Measures Related to the Clarity of Musical Structure

1. *Timing deviation at the ends of phrases: Length of inter-phrase pauses and phrase-final syllables relative to phrase length.* Inter-phrase pauses were relatively greater in infant-directed over infant-absent versions of both playsongs, $t(5) = 2.29$, $p < .03$, and lullabies, $t(3) = 2.48$, $p < .05$. When the onset-to-onset of beats was considered by examining the length of the phrase-final syllable plus inter-phrase pause, this duration also tended to be relatively longer in the infant-directed over infant-absent versions of playsongs, $t(5) = 1.68$, $p < .08$. However, there was no difference for lullabies, $p > .4$. Phrase-final syllables actually tended to be relatively shorter in infant-directed over infant-absent versions of lullabies (3 of the 4 mothers showing this effect), with mean ratios of .264 and .386, respectively. In playsongs, the mean ratio of final syllable to phrase duration was equivalent

for infant-directed and infant-absent versions, and equal to .204. For lullabies, then, it seems that mothers lengthened the inter-phrase pause, but maintained the overall rhythm, compensating for the pause increase by decreasing the duration of the final syllable. For playsongs, mothers simply increased the inter-phrase pause and did not maintain the beat at phrase endings.

2. *Accent structure: Relative intensity of stressed and unstressed syllables.* Although there was a tendency for the relative intensity of stressed to unstressed syllables to be greater in infant-directed over infant-absent versions of both playsongs (4 of 6 mothers, mean ratios of 2.037 and 1.008, respectively) and lullabies (3 of 4 mothers, mean ratios of 1.274 and .909, respectively), there was sufficient variation that these effects did not approach significance.

3. *Accent structure: Relative duration of stressed and unstressed syllables.* The difference between playsongs and lullabies in the relative duration of stressed to unstressed syllables across infant-directed and infant-absent versions approached conventional levels of significance, $t(8) = 1.74, p < .06$. For playsongs, the relative duration of stressed to unstressed syllables was greater for infant-directed over infant-absent versions, $t(5) = 2.09, p < .05$ (mean ratios = 3.42 and 3.12, respectively). For lullabies, there was no significant difference (mean ratios = 1.77 and 1.83, respectively).

4. *Accent structure: Stressed syllable duration relative to phrase length.* The difference between the infant-directed and infant-absent versions in the ratio of the duration of the stressed syllables to the phrase in which they occurred differed significantly between playsongs and lullabies, $t(8) = 2.77, p < .01$. For lullabies, the effect approached conventional levels of significance, $t(3) = 2.20$, $p < .06$, with stressed syllables relatively shorter in the infant-directed versions (means = .26 and .30, respectively). For playsongs there was virtually no difference (means = .21 and .20, respectively).

5. *Accent structure: Vowel duration in stressed syllables.* There were no significant differences across infant-directed versus infant-absent versions for either playsongs or lullabies in the duration of vowels relative to the length of the syllables in which they occurred.

6. *Tempo.* Infant-directed versions were faster for both playsongs, $t(5) = 3.87$, $p < .006$ (mean beat durations = 498 ms and 488 ms, respectively), and lullabies $t(3) = 5.77, p < .005$ (mean beat durations = 559 and 478, respectively). Playsongs and lullabies did not differ significantly in this respect.

7. *Tempo variability.* There were no significant differences in tempo variability across infant-directed versus infant-absent versions for either playsongs or lullabies.

DISCUSSION

The acoustic analyses revealed that infant-directed and infant-absent version of the songs differed in a number of respects. Thus, a number of acoustic parameters have been identified that are potentially associated with adults' perceptions of and infants' preference for infant-directed singing. Further, playsongs and lullabies were similar on some measures, but different on others, providing further evidence that these represent distinct styles of singing.

Consistent with previous reports of infant-directed singing (Trehub et al., in press) and speech (see Fernald, 1991), the infant-directed versions of both playsongs and lullabies were sung at a higher pitch than the infant-absent versions. Given previous reports of infants' preference for higher over lower spoken voices, it is likely that this factor was involved in infants' preference for the infant-directed over infant-absent versions found by Trainor (1996). The use of higher pitch with infants may well be biologically rooted: Morton (1977) presented evidence that across animal species low pitched sounds are used in aggressive and hostile displays whereas higher pitched sounds are used in fright, appeasement, submission, and friendliness. In humans, interactions between caregivers and infants are presumably among the least aggressive of human interactions. On the other hand, simple psychoacoustic considerations might explain infants' preference for high-pitched voices. Higher voices sound less rough because they have fewer harmonics interacting within critical bands (Sundberg, 1994). In any case, adults produce higher pitch when interacting with infants, and infants appear to prefer higher pitched voices.

In Trainor's (1996) study, voice quality differences between the infant-directed and infant-absent versions were salient to adult raters as well as infants. In particular, the infant-directed versions sounded to adults as if the mother were smiling as she sang. Smiling has the effect of shortening the vocal tract, thereby raising the format frequencies (Tartter, 1980; Tartter & Braun, 1994). As indicated in the *Method*, we were unable to measure formants accurately. However, voice quality differences were evident acoustically in that a greater proportion of the energy in the signal was centered at lower frequencies in the infant-directed over infant-absent versions of both playsongs and lullabies. According to Scherer (1986) pleasant emotions lead to faucal and pharyngeal expansion, which in turn leads to the spectral profile characteristic of *wide voice*: relatively more low- than high-frequency energy.

Infant-directed singing has been rated as more emotionally engaging than infant-absent singing (Trehub et al., in press). Two measures associated with increased intensity of emotion, jitter and shimmer (e.g., Bachorowski & Owren, 1995), were both increased in infant-directed over infant-absent playsongs. Jitter was also increased in infant-directed over infant-absent lullabies. Pitch variation is likely also associated with the perception of emotional engagement and is attention-getting. It was greater in infant-directed over infant-absent versions of playsongs, but not lullabies. Thus, the singing of infant-directed playsongs appeared to involve the modification of a greater number of parameters expected to be associated with an emotionally engaging or attention-getting quality than did the singing of lullabies.

Mothers exaggerated cues to the structure of the music when singing both playsongs and lullabies to their infants: for both playsongs and lullabies, inter-phrase pauses were lengthened relative to phrase length. Playsongs and lullabies differed in their rhythmic features, however. The stress patterns of the infant-directed playsongs were exaggerated compared to their infant-absent versions—the relative duration of stressed to unstressed syllables was greater in the infant-directed over infant-absent versions—but this was not the case for the lullabies. The length of stressed syllables compared to the length of the phrase in which they occurred were relatively *shorter* in the infant-directed over infant-absent versions of the lullabies, but did not differ between the playsong pairs. These acoustic measures are consistent with the finding that the playsongs were rated as relatively more rhythmic than their infant-absent pairs, while the lullabies were rated as relatively less rhythmic (Trainor, 1996).

As discussed in the introduction, specific changes in acoustic parameters are associated with the emotions of joy, happiness, and tenderness/affection. However, none of the acoustic profiles of these emotions exactly matches those found for lullabies and playsongs in the present investigation. Playsongs were perhaps closest to the joy profile as outlined in Scherer (1986). In the infant-directed compared to infant-absent versions, pitch, jitter, pitch variability, and intensity variability were all increased. These modifications are consistent with the joy profile and inconsistent with the happiness profile. On the other hand, there was relatively more energy at lower frequencies and the tempo or rate was slower in the infant-directed compared to infant-absent version, which is consistent with the happiness profile but inconsistent with the joy profile. Thus the emotional message of the playsongs appears to be different from any of the primary emotions described acoustically to date.

The few acoustic analyses of affection/tenderness have characterized it as being slower in tempo and with a more regular rhythm (Davitz, 1964). The infant-directed lullabies were indeed slower than their infant-absent pairs and more regular in rhythm. As in joy, the pitch of the infant-directed versions was higher than their infant-absent pairs; however, pitch variability was not greater in the infant-directed versions. The concentration of energy at lower frequencies and the slow tempo were consistent with happiness. Thus, as with playsongs, the acoustic parameters characterizing infant-directed lullabies do not seem to be entirely consistent with the profiles for any of the primary emotions studied in adulthood.

There are a number of possible reasons for the lack of exact correspondence between the acoustic changes found in infant-directed singing and other studies of emotion. First, the vast majority of studies of vocal emotion have examined speech, not singing. The extent to which acoustic correlates of emotional expression differ between speech and song is not known. Second, the few studies of emotion in singing have examined highly trained professional voices.

Sound production in professional and untrained voices may be somewhat different. However, Sundberg (1987) describes a study by Kotlyar and Morozov, where professional singers were asked to sing the same phrases in happy, sorrowful, fearful, angry, or neutral moods. Raters were quite accurate at distinguishing the emotional intent of the singers. Interestingly, when electronically generated versions that maintained the pitch, loudness, and duration characteristics of the originals were presented, raters performed quite well on all the emotions except joy. Thus, it appears that voice quality is extremely important for conveying the emotion of joy in singing.

The lack of exact correspondence between the acoustic features of infant-directed singing and any primary adult emotions may also be caused by the multiple functions of the modifications found in infant-directed singing. These different functions may put some acoustic features in conflict. For example, in both lullabies and playsongs mothers may raise their pitch and slow the tempo in order to clarify the musical structure and attract their infants' attention, even though communicating a soothing message in a lullaby might be better served with a lower pitch and communicating a joyous message in a playsong with a faster tempo. Interestingly, faster speech is rated as less benevolent than speech at a neutral speed (Brown, Strong, & Rencher, 1973), and it might be particularly important with infants not to appear threatening. This interpretation is in line with the findings that the pitch increase was greater for playsongs than lullabies. Overall, playsongs were sung at a higher pitch than lullabies and lullabies were sung more slowly than playsongs.

Finally, it is possible that infant-directed playsongs and lullabies are communicating emotions that do not correspond exactly to the primary emotions, but rather to a complex combination of primary emotions. For example, mothers may be expressing elements of both happiness and joy in playsongs. In lullabies, they may be communicating elements of both happiness and affection/tenderness.

In any case, these results indicate that consistent acoustic modifications in infant-directed singing can be measured. The next step is to look at whether each of these particular modifications have meaning for infants. This could be approached by digitally modifying singing samples to add or remove specific acoustic features and testing infants' reactions to and preferences for versions with or without these changes. Both facial expressions (Ekman, 1989) and vocal expressions (see Frick, 1985) of specific emotions are recognized universally by adults. In this light, it would be of particular interest to examine whether the particular acoustic modifications made by mothers in singing to their infants convey similar meaning to adult raters and infant listeners across cultures.

Authors' Notes

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Science and Engineering Research Board of McMaster University. We are grateful to Terri Lewis for comments on an earlier draft.

REFERENCES

Allen, T.W., Walker, K., Symonds, L., & Marcell, M. (1977). Intrasensory and intersensory perception of temporal sequences during infancy. *Developmental Psychology, 13*, 225–229.

Bachorowski, J.A., & Owren, M.J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional and intensity context. *Psychological Science, 6*, 219–224.

Baken, R.J. (1987). *Clinical measures of speech and voice*. Boston: Allyn and Bacon.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics, 14*, 303–309.

Bever, T.G. (1988). A cognitive theory of emotion and aesthetics in music. *Psychomusicology, 7*, 165–175.

Brown, B.L., Strong, W.J., & Rencher, A.C. (1973). Perceptions of personality from speech: Effects of manipulations of acoustical parameters. *Journal of the Acoustic Society of America, 54*, 29–35.

Caron, A.J., Caron, R.F., & MacLean, D.J. (1988). Infant discrimination of naturalistic emotional expressions: The role of face and voice. *Child Development, 59*, 604–616.

Cooke, D. (1959). *The language of music*. London: Oxford University Press.

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

Cunningham, J.G., & Sterling, R.S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion, 12,* 399–413.

Davis, W.B., Gfeller, K.E., & Thaut, M.H. (1992). *An introduction to music therapy: Theory and practice.* Dubuque, IA: Brown.

Davitz, J.R. (1964). Personality, perceptual, and cognitive correlates of emotional sensitivity. In J.R. Davitz (Ed.), *The communication of emotional meaning* (pp. 57–68). New York: McGraw-Hill.

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

Echols, C.H. (1996). A role for stress in early speech segmentation. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 151–170). Hillsdale, NJ: Erlbaum.

Ekman, P. (1989). The argument and evidence about universals in facial expressions of emotion. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 143–164). Chichester, UK: Wiley.

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

Fernald, A. (1991). Prosody in speech to children: Prelinguistic and linguistic functions. *Annals of Child Development, 8,* 43–80.

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64,* 657–674.

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.

Fonagy, I., & Magdics, K. (1963). Emotional patterns in intonation and music. *Phonetiche Sprachwisson Kommunikationsforsch., 16,* 293–326.

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

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

Gentile, D.A., Pick, A.D., Flom, R.A., & Campos, J.J. (1994). Adults' and preschoolers' perception of emotional meaning in music. Poster presented at the Conference on Human Development, Pittsburgh, PA.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64,* 675–687.

Jusczyk, P.W., Kemler Nelson, D.G., Hirsh-Pasek, K., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24,* 105–109.

Jusczyk, P.W., & Krumhansl, C.L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 627–640.

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.

Langer, S. (1957). *Philosophy in a new key.* 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 Acoustic Society of America, 34,* 922–927.

Mendelson, M.J. (1986). Perception of the temporal pattern of motion in infancy. *Infant Behavior and Development, 9,* 231–243.

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

Morgan, J.L. (1994). Converging measure of speech segmentation in preverbal infants. *Infant Behavior and Development, 17,* 389–403.

Morrongiello, B.A. (1984). Auditory temporal pattern perception in 6- and 12-month-old infants. *Developmental Psychology, 20,* 441–448.

Morton, E.S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *American Naturalist, 111,* 855–869.

Murray, I.R., & Arnott J.L. (1983). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustic Society of America, 93,* 1097–1108.

Papousek, M. (1992). Early ontogeny of vocal communication in parent–infant interactions. In M. Papousek, U. Jurgens, & H. Papousek (Eds.), *Nonverbal vocal communication* (pp. 230–261). Cambridge: Cambridge University Press.

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

Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99,* 143–165.

Sloboda, J.A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19,* 110–120.

Sundberg, J. (1987). *The science of the singing voice.* DeKlab, IL: Northern Illinois University Press.

Sundberg, J. (1994). Perceptual aspects of singing. *Journal of Voice, 8,* 106–122.

Tartter, V.C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics, 27,* 24–27.

Tartter, V.C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *Journal of the Acoustic Society of America, 96,* 2101–2107.

Trainor, L.J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infants Behavior and Development, 19,* 83–92.

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

Trehub, S.E., & Schellenberg, E.G. (1995). Music: Its relevance to infants. *Annals of child development, 11,* 1–24.

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., Trainor, L.J., & Unyk, A.M. (1993). Music and speech processing in the first year of life. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 24, pp. 1–35). San Diego, CA: Academic.

Trehub, S.E., Unyk, A.M., Kamenetsky, S.B., Hill, D.S., Trainor, L.J., Henderson, J.L., & Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology, 33,* 500–507.

Trehub, SE., Unyk, A.M., & Trainor, L.J. (1993a). Adults identify infant-directed music across cultures. *Infant Behavior and Development, 16,* 193–211.

Trehub, S.E., Unyk, A.M., & Trainor, L.J. (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development, 16,* 285–295.

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.

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.

Walker-Andrews, A.S., & Grolnick, W. (1985). Discrimination of vocal expressions by young infants. *Infant Behavior and Development, 6,* 491–498.

05 May 1996; Revised 11 December 1996 ■