Music Perception
Winter 2002, Vol. 20, No. 2, 187–194

Preference for Sensory Consonance in 2- and 4-Month-Old Infants

LAUREL J. TRAINOR, CHRISTINE D. TSANG, &
VIVIAN H. W. CHEUNG
McMaster University

The preferences of 2- and 4-month-old infants for consonant versus dissonant two-tone intervals was tested by using a looking-time preference procedure. Infants of both ages preferred to listen to consonant over dissonant intervals and found it difficult to recover interest after a sequence of dissonant trials. Thus, sensitivity to consonance and dissonance is found before knowledge of scale structure and may be based on the innate structure of the inner ear and the firing characteristics of the auditory nerve. It is likely that consonance perception provides a bootstrap into the task of learning the pitch structure of the musical system to which the infant is exposed.

Received April 10, 2002, accepted September 11, 2002

INFANTS have preferences for certain musical performance characteristics. They prefer to listen to infant-directed over regular singing (Masataka, 1999; Trainor, 1996), higher-pitched over lower-pitched singing (Trainor & Heinmiller, 1998), and singing rendered in a more loving tone of voice (Trehub & Trainor, 1998). At the same time, young infants do not appear to understand the pitch structure of the musical system of their culture (Lynch et al. 1991; Trainor & Trehub, 1992). Adults, even those with no formal musical training, do acquire at least an intuitive understanding, indicating that perceiving through the filter of the pitch structure of a musical system must be learned. It is interesting to ask, then, whether there are precursor abilities that might facilitate or bootstrap the learning of the pitch structure of the music of one's culture. In the study described in this article, we investigated whether infants as young as 2 months old are sensitive to consonance and dissonance.

Address correspondence to Laurel J. Trainor, Department of Psychology, McMaster University, Hamilton, ON, Canada, L8S 4K1. (e-mail: ljt@mcmaster.ca)
ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

Although different musical systems divide the octave into different intervals, and therefore use different sets of notes or scales, the principle of sensory consonance appears to be common across this diversity (Schellenberg & Trehub, 1994). In its simplest definition, two or more tones sounding pleasant together are said to be consonant, whereas two or more tones sounding unpleasant or rough together are said to be dissonant. Consonance and dissonance play a vital role in music in that the ebb and flow of musical tension, which gives rise to emotion and meaning in music, stems in part from the resolution offered by consonance (Meyer, 1956). Of course, in music, context and cultural practice also play a role in the perception of consonance (Cazden, 1980). However, for chords in isolation, the perception of consonance and dissonance appears to arise at relatively peripheral levels of the auditory system (Tramo, Cariani, Delgutte, & Braida, 2001), giving rise to the possibility that sensitivity to sensory consonance and dissonance could develop very early in life.

Plomp proposed that the critical band structure of the basilar membrane in the inner ear can explain the perception of sensory consonance (Plomp & Levelt, 1965). Only simultaneous frequencies within a critical band (about 1/4 of an octave for much of the frequency range) interact in their representation in the ear. Intervals whose tones stand in small-integer frequency ratios (e.g., octaves, 1:2 and perfect fifths, 2:3) sound consonant, whereas more complex frequency ratios (e.g., tritones, 32:45 and minor ninths, 15:32) sound dissonant because in the former cases no or few harmonics or overtones fall within critical bands, whereas in the latter cases, many harmonics or overtones between the two tones fall within critical bands. Tramo et al. (2001) have shown that the consonance/dissonance distinction is also represented in the fine temporal structure of the firing patterns of neurons in the auditory nerve.

Thus evidence suggests that in both the place mechanism (Plomp & Levelt, 1965) and the temporal mechanism (Tramo et al., 2001) consonance and dissonance are encoded in the peripheral auditory system. Given that the peripheral auditory system is relatively mature early in life, one could predict that young infants should be sensitive to this dimension. Indeed this appears to be the case. Infants between 4 and 6 months of age prefer to listen to consonant over dissonant intervals (Trainor & Heinmiller, 1998; Zentner & Kagan, 1998), and infants of 6 months also find two consonant intervals to sound more similar than a consonant and a dissonant interval, even when the latter intervals are more similar in term of size (Schellenberg & Trainor, 1996).

The following experiment investigated whether 2- and 4-month-old infants also have preferences for consonant over dissonant intervals.

Method

PARTICIPANTS

Twenty infants between 8 and 10 weeks old (M = 8 weeks, 6 days; 9 male and 11 female) and 20 infants between 15 and 17 weeks old (M = 16 weeks, 2 days; 9 male and 11 female) were included in the study. All were born within 2 weeks of term, weighted at least 2500 g at birth, and were healthy at the time of testing. An additional 12 infants were excluded for not completing the study because of fussing and/or crying (eleven 8-week-olds and one 16-week-old), and a further 20 (seven 8-week-olds and thirteen 16-week-olds) for yielding low correlations between observers (see below). In the final sample of 40 infants, performance did not differ significantly between males and females.

APPARATUS

The digital sound files were played by a Macintosh IIci computer with an Audiomedia II sound card, through a Denon amplifier (PMA-480R) to a single audiological loudspeaker (GSI). The loudspeaker was located inside a sound-attenuating booth (Industrial Acoustics Co.) on top of a box with a smoked Plexiglas front. Inside the box were lights that, when turned on, illuminated a bull's-eye pattern of alternating black and white concentric circles. The infant sat in a car seat 45 cm from the bull's-eye pattern whose stripes of the pattern were 1.1 cm apart. A video camera (Sony Hi8 Handycam) was located above the box and delivered an image of the infant's face to a monitor (Panasonic CT-1331) located outside the sound-attenuating booth. Two observers watched the monitor and independently recorded their judgments as to when the infant was looking at the bull's-eye pattern by pressing buttons on custom-built button boxes. The lights and the button boxes were connected to the computer via a custom-built interface box to a Strawberry Tree I/O card in the computer.

STIMULI

The stimuli consisted of the same two sets of four simultaneous two-tone intervals used in Trainor and Heinmiller (1998). All tones had piano timbre. The set of consonant intervals consisted of four highly consonant intervals, two perfect fifths (A_3–E_4 [220.0–329.6 Hz] and C_4 to G_4 [261.6–392.0 Hz]) and two octaves (C_4–C_5 [261.6–523.3 Hz] and E_4–E_5 [329.6–659.3 Hz]). The set of dissonant intervals consisted of four highly dissonant intervals, two tritones ($B\flat_3$–E_4 [246.9–329.6 Hz] and F_4–B_4 [349.2–493.9 Hz]) and two minor ninths ($B\flat_3$–B_4 [246.9–493.9 Hz] and E_4 to F_5 [329.6–698.5 Hz]). The consonant and dissonant sets were closely matched in interval size: the tritone is 1 semitone smaller than the perfect fifth and the minor ninth is 1 semitone larger than the octave. Furthermore, with octave equivalence, both the consonant and dissonant sets contain 4 different notes, A C E G and B♭ B E F, respectively, and the range of the two sets is identical (19 semitones). In terms of pitch height, infants would be expected to prefer the dissonant set, if anything, as it is has a slightly higher average pitch, and previous research has established that infants prefer higher to lower pitched music (Trainor & Zacharias, 1998). The two sets differ in that the notes of the consonant set are all members of one key, that of C major, whereas two keys are needed to encompass the four notes of the dissonant set. However, adults tend to hear the consonant set as going between two keys, C major and A minor. In any case, infants of this age do not appear to have knowledge of key structure (Trainor & Trehub, 1992), so it is unlikely that this difference could influence their preferences.

Each consonant and each dissonant interval was presented in a rhythmic pattern (onset-to-onsets of 600, 300, 300, 600 ms; see Figure 1) in order to maintain the infants' interest.

Fig. 1. Sample excerpts from a consonant trial (upper panel) and a dissonant trial (lower panel).

PROCEDURE

Because infants less than 5 months old do not have good control of head movement, and go through a period of not responding to sound location, the head-turn preference procedure used previously with infants older than 5 months (Fernald, 1993; Werker & McLeod, 1989) was modified to a single-speaker setup. In this procedure, infants control how long they listen to the consonant intervals and how long they listen to the dissonant intervals by their looking behavior. Specifically, the infant was seated in the car seat in front of the Plexiglas-fronted box in the sound-attenuating chamber, and presented with eight trials. The two observers watched the infant's face on the monitor outside the sound booth (with no sounds so that they were unaware of what the infant was hearing) and judged by looking at the infant's eyes when the infant was looking at the bull's-eye. Each observer pressed a button when they judged that the infant was ready for a trial. Once both observers had pressed their button (i.e., both buttons were down at the same time), the lights inside the Plexiglas-fronted box began to flash (controlled by the computer), illuminating the bull's-eye and attracting the infant's attention. When the observers judged that the infant was looking at the bull's-eye, they each pressed a second button on their button box to indicate that the infant was ready for the sound. Once both second buttons were depressed, the computer left the lights on, fully illuminating the bull's eye, and the sound for that trial began to play. When the infant was judged to stop looking at the bull's eye, each observer removed their finger from their button. When both observers had removed their fingers for at least 500 ms, the sound and lights were extinguished and the trial ended. The computer kept track of the button-press times of each observer so that interrater reliabilities could be examined. Thus, the looking times on each trial used for data analysis were the longest of the two observers; however, infants were not included in the analyses when there was substantial disagreement between the observers (see Results section for details).

The first and last trials were probe trials in which the vowel /i/, spoken with falling intonation, was repeated every 2400 ms. The purpose of the first trial was to introduce infants to the procedure, and the purpose of the last trial was to examine whether infants were still under procedural control, even if they were bored with the repetition of the stimuli on the preceding trials.

For half of the infants, Trials 2, 3, and 4 were consonant trials, and Trials 5, 6, and 7 were dissonant trials. For the other half of the infants, this was reversed, such that they heard the dissonant trials first and the consonant trials second. Consonant trials consisted of a sequence of the consonant intervals (in the rhythmic pattern described in the Stimulus section; Figure 1) in random order, with each rhythmic pattern separated by 600 ms. Dissonant trials consisted of a similar random order of the dissonant intervals (Figure 1). The computer kept track of looking times across trials for further analysis.

Results

In order for an infant to be included in the final sample, it was required that the correlation between the judged looking times of the two observers across the six consonant and dissonant trials of that infant be greater than .80. In practice, the distribution of correlations tended to be bimodal; the looking behavior of two thirds of the infants (n = 40) was easy to observe and yielded high correlations, but one third of the infants (n = 20) showed erratic looking behavior and yielded low correlations. The correlations for 2-month-old infants included in the final sample averaged .95 (SD = .068, n = 20); the 4-month-olds averaged .95 (SD = .055, n = 20). On the other hand, the correlations for the discarded infants averaged .42 (SD = .033; n = 20).

For each infant, the three consonant trials were averaged and the three dissonant trials were averaged. These looking times were subjected to an analysis of variance with age (2, 4 months) and condition (consonant trials first, dissonant trials first) as between-subject factors and trial type (consonant trials, dissonant trials) as a repeated measure. There was a main effect of age, $F(1, 36) = 17.97$, $p < .0001$, with the younger infants (M = 23.1 s, SD = 12.3) looking overall much longer than the older infants (M = 9.0, SD = 8.1; Figure 2). This result was expected and is consistent with other published reports showing decreasing looking times with increasing age (e.g., Ames, 1988). Looking time reflects an infants' interest in the stimulus. Be-

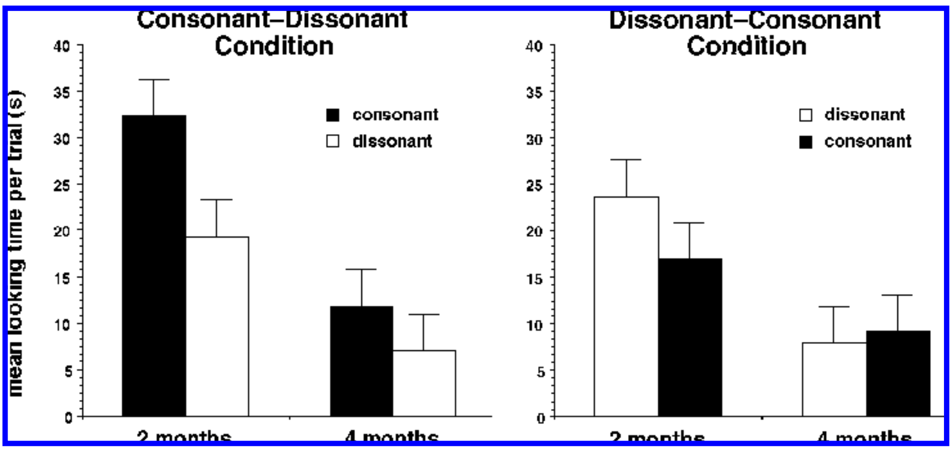

Fig. 2. Mean looking times in seconds to consonant and dissonant trials in 2- and 4-month-old infants. Infants heard either three consonant trials followed by three dissonant trials (left panel) or three dissonant trials followed by three consonant trials (right panel). Note that error bars represent 95% confidence intervals based on the within-subject variability and are calculated according to the formula derived by Loftus and Masson (1994).

cause older infants are more efficient and experienced perceivers, they will process a stimulus of the same complexity more quickly, and thus will become bored with it more quickly, than will 2-month-olds. The important age result from the perspective of the present study is that there were no significant interactions involving age.

Although the effects of trial type and condition were not significant [$F(1, 36) = 1.30$, $p = .26$; $F(1, 36) = .94$, $p = .34$, respectively], there was a significant interaction between condition and trial type, $F(1, 36) = 4.51$, $p < .04$. When infants heard the consonant trials first, they showed increased looking to the consonant trials than to the dissonant trials. However, if they heard the dissonant trials first, they showed low looking times to the dissonant trials and did not recover when presented with the consonant trials. This interpretation was confirmed with two separate analyses of variance, one for each condition. For infants hearing the three consonant trials followed by the three dissonant trials, the effect of trial type was marginally significant, $F(1, 18) = 3.68$, $p = .07$, with infants looking longer to consonant over dissonant trials. However, for infants hearing the three dissonant trials followed by the three consonant trials, the effect of trial type was not significant, $p = .36$.

Habituation effects (decreasing looking times with successive trials as infants become bored) cannot fully explain these results because habituation is not seen when the dissonant trials are presented first. Rather, infants appear to not like the dissonant trials (they show low looking times to dissonant trials whether or not the dissonant trials were heard before or after the consonant trials). They appear to like the consonant trials when they are presented first, but not when they are presented after the dissonant trials. One explanation for this finding is that the dissonant trials "turn the infants off" sufficiently that they do not recover interest during the following consonant trials. Similar asymmetries have been seen in discrimination data. For example, it is much easier for both infants and adults to detect when a dissonant interval is placed in a sequence of consonant intervals than to detect when a consonant interval is placed in a sequence of dissonant intervals (Trainor, 1997). It appears that the dissonant context is difficult to encode and disrupts the encoding of consonant intervals as well.

In sum, the results indicate that infants as young as 2 months of age prefer to listen to consonant over dissonant intervals and that they find it difficult to recover interest after a sequence of dissonant trials.

Discussion

This study establishes that infants as young as 2 months old are sensitive to the dimension of consonance and dissonance, and they prefer to listen to intervals that are consonant. But is early sensitivity to consonance and dissonance specified in the genetic code that informs the development of the peripheral auditory system, or does sensitivity arise through very early experience with real-world sounds, or does sensitivity arise through an interaction between these two processes? The auditory system is functioning by the sixth prenatal month (see Werner & Marean, 1996), and although filtered substantially at the high-frequency end, environmental sounds do reach the fetus. Sounds with pitch, such as music and speech, are prominent in the human environment and have energy mainly at integer multiples of the fundamental. In this overtone structure, the lowest, most prominent and resolvable intervals are also the most consonant, so it is very possible that infants learn about consonance through exposure to such sounds (Terhardt, 1984). Such exposure may, in fact, tune the connections between neurons in the auditory system such that neurons that encode frequencies that are consonantly related become more highly interconnected. On the other hand, the structure of the basilar membrane likely has a strong genetic component, so the role of critical band structure in consonance and dissonance perception may largely arise from a genetic origin. More research will be needed in order to clarify the relative roles of genetic and experiential factors in the perception of consonance and dissonance.

What is clear is that very early in life the auditory system is set up to process an aspect of pitch structure that is critical to the development of musical perception across musical systems. In the majority of musical pitch systems, consonant intervals play an important role; indeed, in most systems, tones an octave apart are functionally similar. Thus, early sensitivity to consonance and dissonance likely sets the stage for learning musical-system-specific pitch structure by providing a bootstrap into important aspects of that structure.[1]

References

Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research, 5,* 69–95.

Cazden, N. (1980). The definition of consonance and dissonance. *International Review of the Aesthetics and Sociology of Music, 2,* 23–168.

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

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Lynch, M. P., Eilers, R. E., Oller, K. D., Urbano, R. C., & Wilson, P. (1991). Influences of acculturation and musical sophistication on perception of musical interval patterns. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 967–975.

1. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. We thank Joanne Leuzzi for her invaluable assistance in the lab and Alan Rosenthal for technical help.

Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology, 35,* 1001–1005.

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

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38,* 548–560.

Schellenberg, E. G., & Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. *Journal of the Acoustical Society of America, 100,* 3321–3328.

Schellenberg, E. G., & Trehub, S. E. (1994). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin and Review, 1,* 191–201.

Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception, 1,* 276–295.

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

Trainor, L. J. (1997). Effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1427–1438.

Trainor, L. J., & Heinmiller, B. J. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development, 21,* 77–88.

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

Trainor, L. J., & Zacharias, C. A. (1998). Infants prefer higher-pitched singing. *Infant Behavior and Development, 21,* 799–805.

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in Western tonal music. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (pp. 92–116). New York: New York Academy of Sciences.

Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and play songs. *Advances in Infancy Research, 12,* 43–77.

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.

Werner, L. A., & Marean, G. C. (1996). *Human auditory development.* Boulder, CO: Westview.

Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development, 21,* 483–492.