Cross-cultural Perspectives on Pitch Memory

Takayuki Nakata

Department of Child Studies
Nagasaki Junshin Catholic University
nakata@n-junshin.ac.jp

Sandra E. Trehub and E. Glenn Schellenberg

Department of Psychology
University of Toronto at Mississauga
sandra.trehub@utoronto.ca g.schellenberg@utoronto.ca

Abstract
We examined pitch memory in young children and adults from Japan and Canada. Listeners heard pairs of 5-second excerpts from familiar soundtracks, one at the original pitch level and the other shifted upward or downward by one semitone. Their task was to identify the original. Adults from both cultures performed above chance levels, but Japanese adults (70% correct) were significantly more accurate than Canadian adults (58% correct). Canadian 5- and 6-year-olds performed at chance levels, which was poorer than the performance of Canadian adults. Japanese children performed well above chance levels (62% correct), but more poorly than Japanese adults. In fact, Japanese children’s performance was midway between that of Japanese and Canadian adults. Different instructional styles in early music education may account for the superiority of Japanese listeners over Canadian listeners.

1. Introduction
Absolute pitch (AP) refers to the ability to identify or reproduce isolated tones in the absence of contextual cues or reference pitches (i.e., naming the corresponding note on a musical scale or singing the named note). Individuals with AP (sometimes called perfect pitch) have excellent long-term memory for musically relevant pitches (e.g., those corresponding to white and black piano keys), but they remember other pitches no better than individuals without AP[1]. In other words, their memory for pitch is enhanced only when they can make use of musical pitch labels. The estimated incidence of AP is one in 10,000, but the incidence is thought to be much greater in Japan, perhaps because of the early onset of music training, along with the practice of using specific verbal labels (do, re, mi) for specific pitches (C, D, E) [2][1]. In the west, the label do generally refers to the tonic note of a musical piece, regardless of its pitch.

Except for individuals who have AP, adults are thought to have relatively poor memory for pitch. Their memory for pitch and other sensory attributes is thought to fade quickly [3]. Pitch memory is also thought to be better in childhood than in adulthood. The latter assumption stems from the early onset of musical training in most individuals with AP. Processing of absolute aspects of pitch is thought to dominate in early childhood, with relative pitch processing dominating later on [4][2]. According to conventional accounts, both modes of processing remain available only for individuals with AP.

There are many reasons to question traditional views of pitch memory and its development. The studies that have reported poor pitch memory on the part of adults have used tasks that make little sense to listeners other than those with AP. For example, they ask listeners to name isolated tones, which are meaningless to most listeners. Moreover, the pitch-naming requirement necessitates musical training. Instead of shedding light on adults’ memory for pitch, those tasks may actually obscure their pitch memory. When adults are given meaningful tasks such as singing the same song on different occasions [5] [6], reproducing songs from popular recordings [7], or recognizing familiar instrumental recordings [8], they show relatively good memory for pitch. Because the production of songs is likely to implicate motor memory as well as pitch memory, recognition tasks provide a better index of pitch memory.

Schellenberg and Trehub [8] presented adults with pairs of instrumental excerpts from familiar television programs. One excerpt was played at the original pitch level, and the other was shifted upward or downward by 1 or 2 semitones. Adults were required to judge whether the first or second excerpt was at the original pitch level. They performed above chance levels on 2-semitone shifts (70% correct) and on 1-semitone shifts (58% correct).

Schellenberg and Trehub’s [8] findings challenge the assumption of adults’ poor pitch memory, but the assumption of children’s superior pitch memory remains unchallenged. Moreover, the implications of different musical practices across cultures are unclear. If the onset and na-
ture of music training in Japan increase the incidence of AP, they may also generate better performance on pitch memory tasks like those of Schellenberg and Trehub [8]. It is possible that cross-cultural differences would be evident in childhood as well.

The purpose of the present investigation was to compare children’s and adults’ memory for the pitch level of familiar music in Japan and Canada. As in the Schellenberg and Trehub [8] study, children and adults were exposed to pairs of excerpts from television programs that they watched regularly. On each trial, they judged which excerpt (first or second) was the original. On the basis of earlier and more intensive music training in Japan than in Canada, we expected Japanese adults to perform more accurately than Canadian adults. For similar reasons, we expected Japanese children to perform better than Canadian children. However, children would have considerably less musical experience than adults. Thus, contrary to prevailing ideas about children’s pitch memory [4] [2], we expected children to perform more poorly than adults in both cultures.

2. Method

2.1. Participants

There were 56 Japanese children and 50 Canadian children who were 5-6 years of age. Six Japanese children had experienced some formal music training (lessons), but no Canadian children had any training. All children were attending school, which included singing along with other activities. For the Japanese children, singing always occurred with piano accompaniment, which ensured that the same songs were produced at the same pitch level on different occasions. The sample of adults included 41 students from a Japanese university and 48 from a Canadian university. Participation was restricted to students who were familiar with the television programs from which the music was extracted. The distribution of music training (i.e., years of music lessons) was bimodal for Japanese students, with peaks at zero (n = 6) and 9 (n = 6). The distribution was skewed for Canadian students, with a mean of 5.1, a median of 3, and a mode of zero. None of the participants reported having absolute pitch.

2.2. Stimuli and Apparatus

The Japanese children’s recordings included instrumental and instrumental plus vocal excerpts from 10 popular TV programs. No instrumental music was available from popular children’s programs in Canada. Thus, all of the Canadian children’s materials were instrumental plus vocal. The musical materials for Japanese adults consisted of 5-second instrumental excerpts from three popular television programs and from nine popular commercial recordings. The musical materials for Canadian adults consisted of 5-second instrumental excerpts from six popular television programs (from [8]). Each recording had multiple instruments, with multiple pure-tone components. The excerpts were saved as CD-quality soundfiles.

The excerpts were shifted in pitch by 1 semitone for Japanese listeners (children and adults) and by 1 or 2 semitones for Canadian listeners (children and adults). The shifting of pitch level was accomplished with Pro Tools (Digidesign) digital-editing software, which is used commonly in professional recording studios. Pitch-shifting did not affect the tempo (speed) or overall sound quality of instrumental excerpts. For each musical selection, the “incorrect” excerpt was always shifted in one direction (upward for three, downward for three) to eliminate the option of selecting the middle pitch level and to ensure that “correct” and “incorrect” excerpts were presented equally often. The direction of pitch shifts was reversed for half of the sample. Pitch shifts involved multiplying (for upward shifts) or dividing (for downward shifts) all frequencies in an excerpt by a factor of 1.12 for 2-semitone shifts and 1.06 for 1-semitone shifts.

To eliminate potential cues from the electronic manipulation, we also shifted the pitch level of the “correct” excerpts. The original excerpts were shifted upward and downward by 1 semitone (all frequencies multiplied and subsequently divided by 1.06) in the 2-semitone condition; the upward and downward shifts were half a semitone (frequencies multiplied and divided by 1.03) in the 1-semitone condition. For Canadian participants (children and adults) and Japanese adults, monaural excerpts were presented binaurally over lightweight headphones. For Japanese children, monaural excerpts were presented binaurally by speakers connected to a computer. Studies took place in a quiet room for Japanese participants and in a sound-attenuating booth for Canadian participants.

2.3. Procedure

Adults were tested on six musical selections. The musical selections were chosen on the basis of each participant’s reported familiarity with the TV programs or popular songs from which the excerpts were drawn. Japanese adults were tested on 1-semitone shifts in one test session. Canadian adults were tested in two test sessions on different days no more than one week apart. In one session, Canadian adults were tested on 1-semitone shifts; in the other (counterbalanced order), they were tested on 2-semitone shifts. Each of the adult sessions consisted of five blocks of six trials. Each block had one trial for each excerpt, with trials presented in random order. The first block served as a practice block. On each trial, listeners heard one version of a 5-second excerpt at the original pitch level and another at the altered (upward or downward) pitch, with the two excerpts separated by 2 seconds. Order (original-altered or altered-original) was counterbalanced. Participants were told that they would hear two versions of the same theme song on each trial,
with one version at the correct pitch and the other version shifted higher or lower. Their task was to identify the excerpt (first or second) at the correct pitch level (i.e., the one that they usually heard). They received no feedback for correct or incorrect responses. Adults also completed a brief questionnaire about their musical background, and they provided estimates of the number of times they had watched each program. Japanese and Canadian children were tested in one session. Half of the Canadian children were tested on 1-semitone shifts, the other half on 2-semitone shifts. The children were tested on music from 4 TV programs that they watched regularly. In other respects, the procedure for children was the same as that for adults.

3. Results

Japanese children’s performance (62% correct) exceeded chance levels (50% correct), t(55) = 4.90, p < .001, as did the performance of Japanese adults (70% correct), t(41) = 4.00, p < .001. As predicted, the performance of Japanese children was poorer than that of adults, t(95) = 1.87, p < .05 (one-tailed). The performance of Japanese adults was highly consistent, with only 5 of 41 performing below 50% correct (binomial test, p < .001). Years of musical training were unrelated to Japanese adults’ performance accuracy, r = .128, p = .213 (one-tailed).

Because we were interested in comparisons between Canadian and Japanese listeners, our presentation of Canadian data is restricted to the 1-semitone shifts. Unlike the Japanese children, Canadian children performed at chance levels. The performance of Canadian adults (58% correct) exceeded chance levels, t(47) = 4.00, p < .001, but it was well below the performance of Japanese adults, t(87) = -3.73, p < .001 (one-tailed).

4. Discussion

Our findings, along with those of Schellenberg and Trehub [8], indicate that good pitch memory is not dependent on formal musical training. When the musical material is familiar, as in the case of television soundtracks or popular tunes, listeners exhibit good memory for pitch level. In line with our predictions, Japanese children and adults remembered the pitch level of familiar recordings more accurately than did Canadian children and adults. Because cross-cultural differences are apparent in early childhood, before most children begin taking music lessons, we cannot attribute Japanese listeners’ superiority to formal musical training. It is possible, indeed likely, that early educational practices are responsible for the cross-cultural differences. Piano accompaniment for children’s singing—a common situation in Japanese kindergartens but not in Canadian kindergartens—provides repeated exposure to the same musical material at the same pitch level, a practice that may enhance children’s attention to pitch level and, consequently, their memory for pitch.

Our data allow us to reject the notion of enhanced pitch memory in early life. Instead, they are consistent with the hypothesized improvement in memory between childhood and adulthood. As expected, Canadian adults performed better than Canadian children, and Japanese adults (70% correct) performed better than Japanese children (62% correct). However, comparisons of university students with 5- to 6-year-old children are likely to place children at a disadvantage. Maintaining attention on any task, especially the challenging two-interval forced-choice task, is likely to be more problematic for children than for adults. Nevertheless, Japanese children’s ability to perform above chance levels indicates that the task was feasible for children of that age level. Although Canadian children performed at chance levels on the 1-semitone shifts, they performed well above chance levels on larger (2-semitone) shifts, which indicates that the magnitude of the pitch change was the major influence on their performance. Further research is necessary to determine which aspects of musical exposure enhance absolute pitch processing in childhood and adulthood. It would also be of interest to determine whether child and adult listeners retain other absolute features of music such as its tempo (speed) and timbre (sound quality).

5. References

2003.