Acoust. Sci. & Tech. **29**, 5 (2008)

PAPER

Functional differences between tonotopic and periodic information in recognition of transposed melodies: How do local cues affect global features?

Toshie Matsui*[,1] and Minoru Tsuzaki[2]

[1]*Graduate School of Music, Kyoto City University of Arts*
[2]*Faculty of Music, Kyoto City University of Arts*

(*Received 25 August 2007, Accepted for publication 21 February 2008*)

Abstract: To recognize a transposed melody, two properties can be hypothesized to function as a significant invariant feature: a melodic contour, an up-down movement; and melodic intervals, the distances between consecutive pitches. Both properties are realized by changing the frequency of notes. Frequency changes are coded in two different ways in the early stages of the auditory process: tonotopicity and periodicity. In the first experiment, to investigate how differently transposed melodies are recognized depending on the two types of cue, two stimulus types were used: band noise (BN) providing mainly the tonotopic cue, and rippled noise (RN) providing the periodic cue. The results indicated that the contour invariance could be extracted from both cues. It was also indicated that the interval size information was contaminated in the contour information. A second experiment tested functional differences between the BN and RN stimuli in providing tonal contexts. The results showed that only the RN stimulus produced a "tonality" effect, which was reflected by an acute sense of "equality." The results verify that the periodic cue provides more accurate information for the melodic interval recognition than the tonotopic cue. The current study reaffirmed that local cues individually affected global features, that is, melodic contour and tonality.

Keywords: Tonotopic cue, Periodic cue, Melodic contour, Melodic interval, Tonality

PACS number: 43.66.Hg [doi:10.1250/ast.29.309]

1. INTRODUCTION

One of the important characteristics of melodies is that the identity of a melody can be preserved even when a melody is transposed. Theoretically, what is preserved through transposition is the relationship between the notes composing the melody. This is one reason why pitch intervals, such as an octave, a perfect fifth, and a major third, have been considered basic elements in Western tonal music. However, it can be noted that a melodic contour, an up-down pattern of consecutive pitches, can function as a more primitive feature for recognizing a melodic sequence. To explain what is the invariant information in recognizing a transposed melody, several previous studies have provided some empirical data. For example, Dowling and Fujitani argued that a melodic contour is the most important feature for recognizing a transposed melody. They also argued that precise judgment about pitch intervals acquired through a certain degree of musical training was a secondary crucial cue [1].

In their discussion of melodic indentification, Dowling and Fujitani [1] did not pay any specific attention to the two different cues provided by frequency changes, i.e., the tonotopic and periodic cues. The main objective of the current study was to investigate how differently these two cues work in constructing melodic contours or pitch intervals.

Technically, melodic transposition can be achieved by multiplying a certain constant value to the fundamental frequency of each tone composing a melody. Or it can be achieved by moving the trajectories of fundamental frequencies parallel on a logarithmic scale, which is assumed to be closer to the human perceptual scale. The transposition operation of a form of function T can be written as:

$$T(\log F_k) = \log F_k + a, \tag{1}$$

where $\log F_k$ is the k-th fundamental frequency in the original, melodic sequence, a is a constant defined by transposing intervals from the original melody. This

*e-mail: m04905@kcua.ac.jp

operation preserves the ratio between the fundamental frequency of tones, while it changes their absolute value.

When a fundamental frequency changes, two aspects of the representation for the tone will change in the early stage of the auditory process. One is the tonotopic aspect, the other is the periodic aspect. Tonotopy corresponds to the difference in the positions at which the basilar membrane shows its prominant oscillation, and is represented as an excitation pattern. Physiologists have found that the tonotopic organization is "preserved" across all of the auditory pathways up to the auditory cortex. Periodicity corresponds to the time interval information observed in neural firings in auditory nerve fibers. It preserves the periodicity of the original waveform.

This periodicity information can serve as a physiological basis for the musical pitch relations because the hierarchy of pitch intervals is structured depending upon the proportion of the common periods. For example, an octave relation has common periods every other cycle [2]. There is no direct evidence for octave equivalence in terms of the tonotopic cue.

On the other hand, it has been reported that the sensation of an up-down pattern depending on periodic information is unstable or ambiguous [3]. The periodic information has cyclic characteristics in its nature. For example, a musical note, which is usually a harmonic complex tone, shares half of its harmonic components and the pitch name with an octave higher note which is twelve steps upwards on the chromatic scale. Because of this circular nature, the periodic information could not provide an appropriate cue for the auditory system to extract the melodic contour, which is the sensation of the up-down trajectory on a one-dimensional scale.

Accordingly, it can be hypothesized that the melodic contour is mainly established from tonotopic information, and that melodic intervals are mainly established by periodic information. Experiment I was designed to test this hypothesis. To exclusively dominate one of the two cues, i.e., tonotopic, or periodic, in the perceptual process of melodies, the recognition of transposed melodies was examined using two types of stimulus to compose notes: band noise (BN) stimuli mainly providing tonotopic information and rippled noise (RN) stimuli mainly providing periodic information.

2. EXPERIMENT I

The purpose of Experiment I was to investigate whether there was any functional difference between the two types of cues in conveying each melodic contour and melodic interval information used in the recognition of a transposed melody.

As mentioned in Chap. 1, if the tonotopic cue conveys melodic contour information while the periodic cue conveys melodic interval information, a melody composed of BN is difficult to discriminate from a melody with a similar melodic contour even when there was a slight mismatch in intervals. In contrast, a melody composed of RN could be confused with a melody with the same melodic intervals, even when the contours are dissimilar.

2.1. Method

2.1.1. Participants

Four volunteers (including the first author) from Kyoto City University of Arts participated in Experiment I. They were highly motivated because they were also graduate students majoring in the psychology of hearing. No listener had any significant hearing deficit.

2.1.2. Stimuli

Two types of stimulus were used to compose melodic sequences for the experiment: BN and RN. In the BN stimuli, the center frequencies were controlled corresponding to the fundamental frequencies of the target melodies. BN stimuli were generated by filtering Gaussian noise through a fourth-order Butterworth filter with a 1/3 octave passband (the roll-off was 30 dB per 1/3 octave). RN stimuli were generated by adding a copy of Gaussian noise to the original with delay [4–6]. The noise was added to itself with a delay corresponding to the inverse of the target fundamental frequency, and both were then filtered through a bandpass filter. The cutoff frequencies were 100 Hz and 2 kHz, and the roll-off frequencies were 0 Hz and 3.6 kHz. For both types of noise, four "original" samples of Gaussian noise were used to avoid artifacts caused by a specific noise seed. The spectra of the two types of noise are shown in Fig. 1.

Melodies were composed of five notes using either BN or RN. The inter-onset interval (IOI) was fixed at 250 ms. Considering the possibility that tonotopic and periodic information have different threshold functions of duration, two conditions of note duration were prepared, i.e., 50 ms and 200 ms, for both types of stimuli (BN and RN) [7].

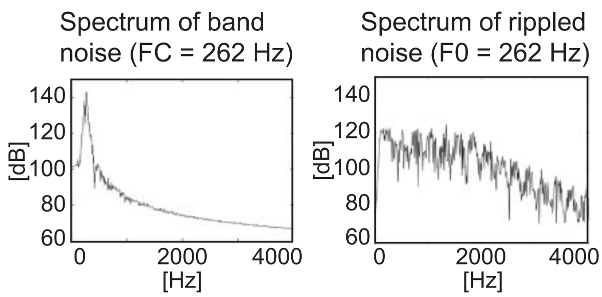

Fig. 1 Power spectra of BN (left panel) and RN (right panel). Center frequency (F_C) for BN and fundamental frequency (F_0) for RN were C4 (262 Hz).

The task was to recognize the exactly transposed version of the original melody. To avoid artifacts caused by using only one specific melodic contour, four standard patterns of melodic contour were prepared, as represented by [+ + + +], [− − − −], [+ + − −], and [− − + +], where the plus [+] sign indicates an upward shift from one note to the next, and the minus [−] sign indicates a downward shift. The exactly transposed melodies (the correct alternatives) were generated by simply multiplying the base fundamental frequencies of the original melody with a certain constant value, which will be described in Sect. 2.1.3. As a result, the transposed melodies had melodic contours identical to their originals. Altered transposed melodies (the false alternatives; called "lure melodies" in the following section) were also based on these four standard patterns of melodic contour.

To specify the cue used for the recognition of transposed melodies, three types of lure melodies were prepared. (A) The first type is called the contour-preserved lure (CP lure) where the melodic contour is almost identical to that of the exactly transposed melody while some of the melodic intervals between adjacent tones are modified. The deviation of intervals in this type of lure melody from the original one is a semitone (100 cents) in half of the lure melodies, and a whole tone (200 cents) in the remaining half. (B) The second type is called the contour-destroyed, pitch-class-preserved lure (CD-PP lure), where the second or fourth tone in the exactly transposed melody is shifted one octave higher or lower. It is assumed that the "compound interval" is equivalent to the original simple interval. The "compound interval" is defined as an interval that is one (or more) octave(s) wider than the original either by replacing the lower note with a one (or more) octave(s) lower note or by replacing the higher note with a one (or more) octave(s) higher note. In either case, the pitch class of the replacing note is the same as the orginal. For example, the interval between C4 and G5 is a perfect 13th; however, following musical theory, it can be regarded as a perfect 5th. To make both the CP and CD-PP lure melodies, the second or fourth notes are changed. (C) The third type is called the contour and pitch-class-preserved lure (CP-PP lure) where one tone in the exactly transposed melody is made one octave higher or lower in a way that the melodic contour is preserved. The frequency of the altered note being shifted to exactly one octave higher or lower than the original is common with the second condition. The critical difference was that the alteration does not destroy the direction of frequency movement between consecutive tones in this type, whereas it does in the second type. For example, the third note was altered to one octave higher in the case of the standard melody pattern [+ + − −]. In the case of [+ + + +], the first note was altered to one octave lower or the last note was altered to one octave higher in the range from C3 to C6. Because of this strict limitation in preparing lure melodies, only one lure can be made from a standard melody pattern, while two lures can be produced from a standard melody pattern under the CP and CD-PP conditions. These three lure types are shown in Fig. 2.

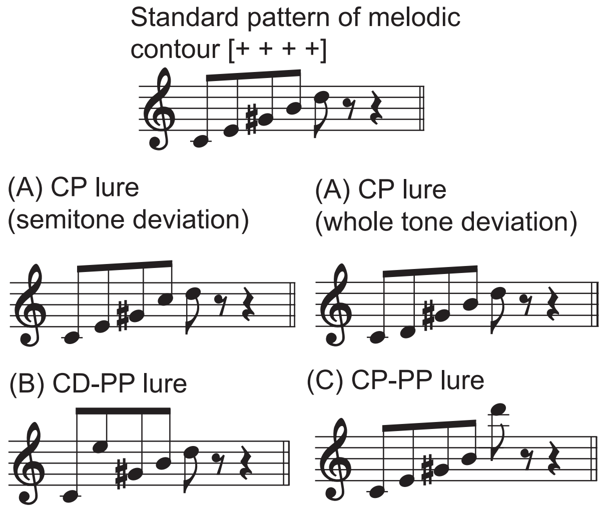

Fig. 2 Examples of three lure melody conditions: (A) contour-preserved lure (CP lure), (B) contour-destroyed, pitch-class-preserved lure (CD-PP lure), and (C) contour and pitch-class-preserved lure (CP-PP lure). Example (A) includes two conditions of deviation, a semitone-deviated lure in the case of changing the fourth note, and whole tone (two semitones)-deviated lure in the case of changing the second note.

2.1.3. Procedure

The recognition of the transposed melody was tested adopting the two alternative forced choice (2AFC) paradigm. In each trial, three melodic pieces were presented. The time interval between melodies was 1,750 ms. The first one was the standard stimulus, i.e., the original melody. The second and the third pieces were either the exact transposition or the lure. Participants were asked to choose which they thought to be the exact transposition, the second or the third. The exactly transposed melody and the lure melody were presented in random order. All the melodies were transposed randomly within the range from C3 (131 Hz) to C6 (1,048 Hz). Transposing intervals were not continuous but discrete (semitone steps).

Participants were allowed to perform the task at their own pace. After each response, feedback indicating which was the correct selection was presented visually. The next trial started automatically a short while after the feedback.

Each participant responded to 320 trials in total: that is, eight standard melodies (two melodies for each standard melody pattern), five lure melody patterns (two for CP lure, two for CD-PP lure, and one for CP-PP lure), two types of stimulus (BN and RN), two duration conditions (50 ms and 200 ms), and their counterbalanced order between the exact transposition and the lure. The average presentation level of stimuli was 70 dB SPL. The sound pressure level was randomly varied within ±10 dB. Stimuli were generated offline using MATLAB (The MathWorks Inc.) on a computer (Apple, PowerMac G5). All the experiments were performed with a GUI on the Apple PowerMac G5. The stimuli were presented using software that controlls a DSP system (Kyma with Capybara 320, Symbolic Sound Corporation). The stimuli were diotically presented through a set of Sennheiser HD 600 headphones. The experiment took place in a sound-insulated booth at the Kyoto City University of the Arts.

2.2. Result

Three-way ANOVA (lure melody condition × duration × stimulus type [BN vs RN]) was performed on the correct response rate pooled over all participants. The standard melody pattern was excluded from factors because it was a random factor. The main effects of the lure melody condition and the stimulus type were significant [$F(2,36) = 144.6$, $p < 0.0001$, $F(1,36) = 30.15$, $p < 0.0001$, respectively]. The interaction between the lure melody condition and the stimulus type was also significant [$F(2,36) = 37.43$, $p < 0.0001$]. Since the factor of duration was not significant and did not interact with the other factors, it was pooled, and two-way ANOVA (lure melody condition × stimulus type) was performed again. As a result, the main effects of the lure melody condition and stimulus type were again found to be significant [$F(2,42) = 140.3$, $p < 0.0001$. $F(1,42) = 29.26$, $p < 0.0001$, respectively]. The interaction between the lure melody condition and stimulus type was also significant [$F(2,42) = 36.32$, $p < 0.0001$].

Correct response rates for each lure melody condition and each stimulus type are shown in Fig. 3. As the task was 2AFC, the chance level was 0.5. When the note was shifted one octave higher or lower, the performance was almost perfect with the RN stimulus type, whether the original contour was destroyed (in lure (B)) or preserved (in lure (C)). This indicates that the compound intervals are not perceived to be the same as their corresponding counterparts in the sense of distance between tones even if they have common elements of periodicity. The concept of octave equivalence is unlikely to hold when compound intervals are applied for the recognition of transposed melody.

As illustrated in Fig. 3, when the target note shifted by a semitone or a whole tone (CP lure), the performance was worse than that under any other condition. As indicated by the significant interaction between the lure condition and stimulus type factors in the ANOVA, the correct response rate of the BN stimulus was significantly worse than that of the RN stimulus under the CP lure condition. The difference between stimulus types indicates that the accuracy of melodic contour information might be affected by the type of element stimulus that compose the melodies. However, because the correct response rate under all the conditions was above the chance level (0.5), it was likely that melodic interval information was not entirely deprived, but that the accuracy of judgment was decreased. Under the lure melody condition of only the contour being preserved, there were two degrees of deviation (a semitone, or a whole tone) from the exactly transposed melody. To observe the effect of this degree of deviation and the stimulus type, a two-way ANOVA (stimulus type × deviation) was performed for data obtained with the lure in which only the contour was preserved (lure (A)). The main effects of stimulus type and deviation were significant [$F(1,28) = 54.09$, $p < 0.0001$. $F(1,28) = 33.61$, $p < 0.0001$, respectively]. The performance with the RN stimulus was better than that with the BN stimulus. The correct response rates under each condition are shown in Fig. 4. Under the BN condition, the correct response rate was at the chance level

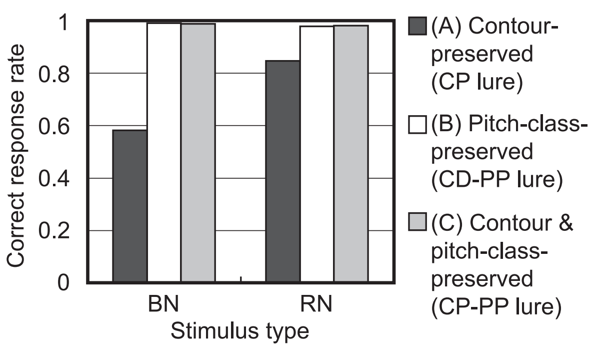

Fig. 3 Correct response rates for each stimulus type [BN vs RN] and lure melody condition: CP lure condition (A) shown as dark grey bars, CD-PP lure condition (B) shown as white bars, and CP-PP lure condition (C) shown as light grey bars.

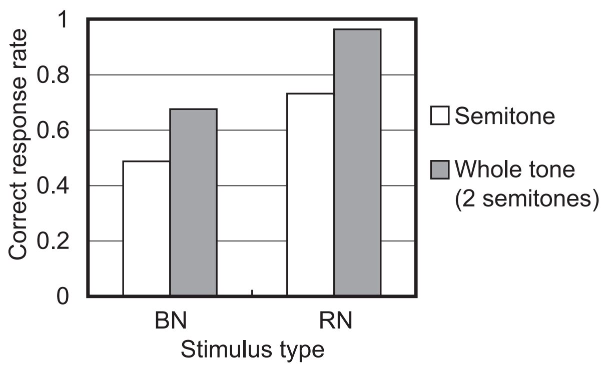

Fig. 4 Correct response rates depending on the degree of interval deviation (semitone vs whole tone) under CP lure condition (A) as a function of stimulus type (BN vs RN). Sizes of deviated intervals are indicated by color of bars: white bars correspond to a semitone, and grey bars correspond to a whole tone (two semitones).

when the deviation was a semitone [$\chi^2(1) = 0.100$, $p = 0.7518$]. This implies that the deviation of a semitone is insufficient as a cue for discrimination.

2.3. Discussion

The results of Experiment I did not support the hypothesis (a) that melodic contour information is mainly conveyed by the tonotopic cue, or (b) that melodic interval information is mainly conveyed by the periodic cue. If the tonotopic cue had exclusively conveyed the contour information, and if octave equivalence had strongly affected melodic interval perception, the correct response rate under the contour-preserved lure melody conditions (CP lure, and CP-PP lure) should have been lower with BN stimuli due to the confusion between the target and lure that could be caused by their sharing the same contour. On the other hand, if the periodic cue had exclusively provided the pitch class without providing any information on pitch trajectories, the correct response rate under the pitch-class-preserved lure melody conditions (CD-PP lure, and CP-PP lure) should have been lower with RN stimuli due to the confusion between the target and lure that could be caused by their sharing exactly the same pitch class.

It seems that intervals (including octave compounds) can be conveyed by both tonotopic and periodic information. However, a difference was observed in their accuracy. It was observed that a semitone step was indiscriminable only in the case of a tonotopic cue, as indicated by the leftmost bar in Fig. 4. Figure 4 also indicates that the deviation size matters even when the task is discrimination between melodies sharing the same contour. Following the definition of contour by Dowling and Fujitani [1], the concept of the melodic contour should have been a series of simple discrete codes indicating the direction of pitch change, such as [$+ + - -$] and [$+ - + -$]. If the listeners encoded the stimlus with this discrete coding, the interval size (semitone vs whole tone) should produce no significant difference under the BN condition. The melodic contour that is defined as the up-down trajectory is unlikely to be a separate feature from the melodic intervals.

Dowling and Fujitani [1] used atonal melodies [8]; we also used atonal melodies in Experiment I. In the previous studies in which the role of the melodic interval in the recognition of transposed melody was discussed, melodic intervals were considered to be effective codes for salient long-term memory of melodies [9–12]. Moreover, in the previous study of the interaction between transposed melody recognition and tonality, it was shown that well-known tonal melodies could be transposed and replayed easily even by participants without specific musical training [13]. Some experimental data also suggested that tonality contributed to good recognition of a melody [14,15]. The judgment process in Experiment I might be different from that in the transposed tonal melody recognition in the previous studies. Attneave and Olsen demonstrated difficulty in replaying a transposed melody in a different register from the original melody when the original melody was atonal [13]. The structure of the mental representation for melodic intervals in tonal contexts is different from that in atonal contexts [16]. Pitch interval discrimination was more accurate in tonal contexts [17]. The dispersion of pitch interval tuning was significantly narrow in tonal contexts [18]. These findings provide evidence that the perception of pitch intervals in tonal contexts is different from that in atonal contexts.

On the basis of the results of the previous studies and the current study, interactions between two factors, i.e., melodic interval perception and tonal contexts, will be expected in transposed melody recognition. For further discussion of melodic interval perception, it is necessary to investigate the effects of tonal contexts on the perception of melodic intervals, and to observe whether the effect of tonal contexts depends on the differences in the dominant cue of each note, i.e., the tonotopic and periodic cues.

3. EXPERIMENT II

As discussed concerning the results of Experiment I, it is worth investigating the functional difference between the tonotopic and periodic cues in providing the tonal context. In Experiment II, we investigated how the perception of the melodic interval would differ depending on the tonal context provided by either the tonotopic or periodic cue. As mentioned for Experiment I, the effect of tonal context on melodic interval perception has been tested in several previous studies. Among those studies, Tsuzaki performed experiments of melodic interval identification in the tonal and atonal contexts, and reported quantitatively how the tonal context affected the perceptual representation of melodic intervals [19]. He argued that the “equality” aspect of judging melodic intervals was enhanced in the tonal context. In the experiments, the context was manipulated by changing the preceding scales: diatonic scale, chromatic scale, or no-scale condition. In the case of a preceding diatonic scale, the sense of equality was enhanced.

Tsuzaki used harmonic complex tones consisting of six successive components in making the preceding scales as well as the standard and comparison melodic intervals. Accordingly, the context was essentially controlled by the relative relationship between the fundamental frequencies of the scales and target intervals. As demonstrated in Experiment I, even if the same frequency relationship existed, the perceptual effect could be different depending on whether it was provided by the center frequency of BN, or by the reciprocal of the delay time of RN. Thus, a question arises: is there any difference in the tonal effect between the melodic interval provided by the tonotopic cue and that provided by the periodic cue? The purpose of Experiment II was to investigate such a functional difference between tonotopic and periodic cues in the judgment of melodic pitch intervals.

3.1. Method

3.1.1. Participants

Six volunteers from Kyoto City University of Arts participated in Experiment II. No listener had any significant hearing deficit.

3.1.2. Stimuli

As in Experiment I, two types of noise were used for the experiment: BN and RN. Each stimulus was generated in the same way as in Experiment I. However, the range of the center frequency for BN and the delay time for RN were different. The duration of each tone was 105 ms, which was also different from that in Experiment I.

In each trial, participants listened to a sequence consisting of a preceding scale, a standard interval, and a comparison interval. The preceding scale was a major diatonic scale, which began with C4 (262 Hz) and proceeded upward for one octave. Thus, the pitch in the diatonic scale contained C4, D4, E4, F4, G4, A4, B4, and C5. The intervals between adjacent tones were 200, 200, 100, 200, 200, 200, and 100 cents.

The standard intervals were composed of two successive tones. The first tone was either the "leading tone" (B4, 494.6 Hz) or "tonic" (C5, 524 Hz) of the tonal context provided by the preceding tonal scale. The second tone was higher than the first one by 100, 150, or 200 cents. Hence, there were six standard intervals. The comparison intervals also consisted of two successive tones. The first tones always had the same frequency as the second tone of the standard interval. Figure 5 indicates the procedure schematically. The comparison intervals were equal to, smaller than, or larger than each standard interval by 20 cents. To help the listeners to answer with confidence as well as to keep their motivation reasonably high, comparison stimuli that were larger or smaller than standard intervals by 40 and 60 cents were also included in the test set. These stimuli were excluded from the analysis.

All tones had a duration of 105 ms, including 5 ms linear on and off ramps. The inter-onset intervals between tones were 200 ms in the preceding scale, standard intervals, and comparison intervals. A 400 ms duration of silence was set prior to the standard interval and the comparison interval. The average presentation level of stimuli was 70 dB SPL. The overall level was varied across each interval presentation over a 10 dB range to restrain listeners from using sound level as a cue for discrimination.

3.1.3. Procedure

The experimental procedure was mostly identical to that in the previous study [19]. Six listeners were asked to

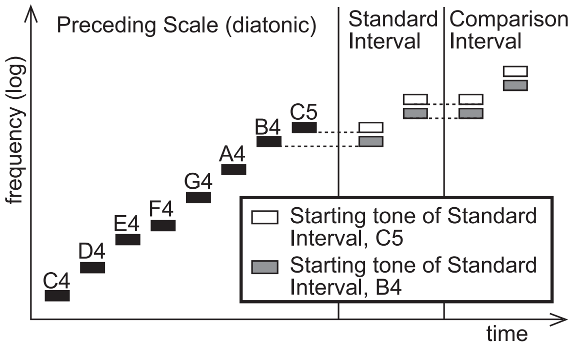

Fig. 5 Schematic representation of sequence of stimuli for Experiment II. The first tone of each standard interval was either B4 (494.6 Hz) or C5 (524 Hz). Second tone of each standard interval was higher than the first by 100, 150, or 200 cents. The first tones of each comparison interval always had the same frequency as the second tone of the preceding standard interval. Comparison intervals were equal to, or smaller, or larger than each standard interval by 20 cents.

discriminate the comparison interval (C) from the standard interval (S) in terms of three response categories, i.e., "smaller ($C < S$)," "equal ($C = S$)," and "larger ($C > S$)." Listeners were allowed to listen to the stimulus repeatedly before they answered. Visual feedback was provided after each answer. All the experiments comprised of seventy-two combinations: two starting tones (B4 and C5), three standard intervals (100, 150, 200 cents), three comparison intervals (-20, 0, $+20$ cents), two types of noise (RN and BN), and their counterbalanced order. Those combinations were presented ten times in total to each listener. Stimuli were presented in a random order sets grouped by stimulus type, BN, or RN. The order of sets was also randomized for each listener. The apparatus and environment for Experiment II were identical to those for Experiment I.

3.2. Results

Figure 6 shows the results for the standard intervals and the starting notes of the standard interval pooled over all participants. The top panels (a) and (b) show the results for the condition that the standard interval was 100 cents starting at C5. Under these conditions, both the standard and comparison intervals are "atonal" because neither of them fits in the preceding diatonic scale. The lower panels (c) and (d) show the result for the standard interval of 200 cents starting at C5. The intervals under these conditions can be regarded to be "tonal" because both are in accord with the diatonic context given by the preceding scale. The left panels (a) and (c) show the results by the BN stimulus condition, while the right panels (b) and (d) show those by the RN stimulus. The abscissa on each panel corresponds to the condition of the comparison interval, while the ordinate represents the response ratio for each response category.

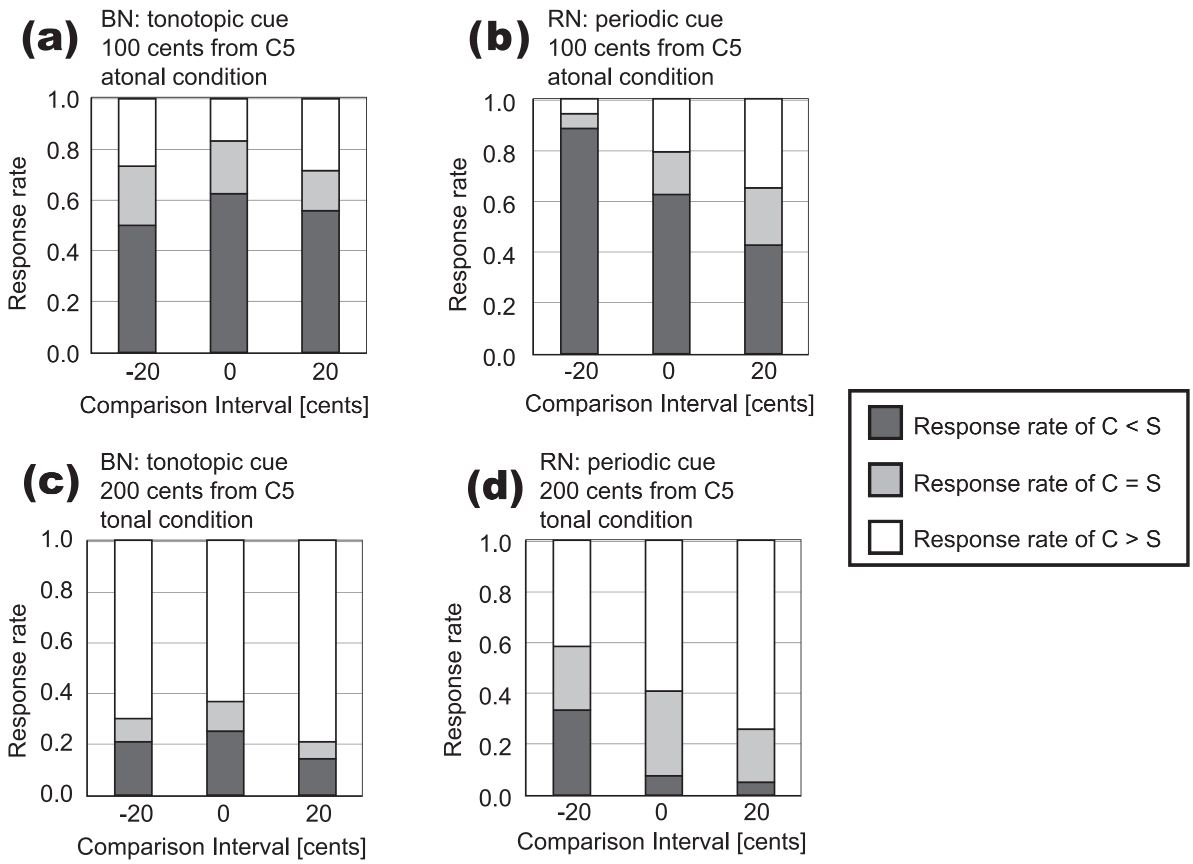

Fig. 6 Results for atonal condition where standard interval was 100 cents starting at C5 (524 Hz), as shown in upper panels (a) and (b). Results for tonal condition where standard interval was 200 cents starting at C5, as shown in lower panels (c) and (d). Left panels (a) and (c) are for the BN condition while right panels (b) and (d) are for the RN condition. Abscissa of each graph corresponds to comparison interval conditions of −20, 0, and 20 cents. Each bar indicates response rate for each response category: response rate where comparison interval was judged to be smaller than standard interval ("$C < S$") shown at bottom as dark grey bars; response rate of "$C = S$" shown in middle as light grey bars; and "$C > S$" shown at top as white bars.

First, these results indicate that there was a tendency for the response rate of the "$C < S$" category to be higher under the 100-cent standard interval condition, and that the response rate of the "$C > S$" category increased under the 200-cent standard interval condition. Second, as shown in panel (d) of Fig. 6, the result for the RN stimuli in the 200-cent standard interval starting at C5 showed a high response rate in the "$C = S$" category for the 0-cent comparison interval. Such a high rate was not observed under any other condition. Since all the intervals under this condition fit in the tonal context, it can be assumed that enhancement of the sense of equality was provided by the tonal context [19].

3.3. Analysis Model

The mental representation for each melodic pitch interval can be estimated in the form of a statistical model based on the results shown in Fig. 6. Generally, when a paired comparison with standard stimulus S and comparison stimulus C is performed, it is assumed that a participant subtracts an internal value of C from an internal value of S on the listener's psychological continuum and then judges whether C is greater or smaller than S according to the sign of the subtraction result. The value of a stimulus on the psychological continuum does not uniquely correspond to the physical value of a stimulus, owing to factors such as internal noise. It is reasonable to assume that the psychological value of a stimulus fluctuates as a stochastic process, and that it can be represented as a probability distribution. On this assumption, an observation of the difference between C and S can be considered to reveal the difference between a distribution for standard stimulus S (ψ_S) and a distribution for comparison stimulus C (ψ_C) on a psychological continuum. This process is represented by the following equation, where dCS indicates the difference of C from S on a psychological continuum:

$$dCS = \psi_C - \psi_S. \qquad (2)$$

When the observation is repeated, the probability density function, ψ_{dCS}, can be considered to give an approximation of the distribution of dCS. It is generally assumed to follow a normal distribution. This is expressed by the following

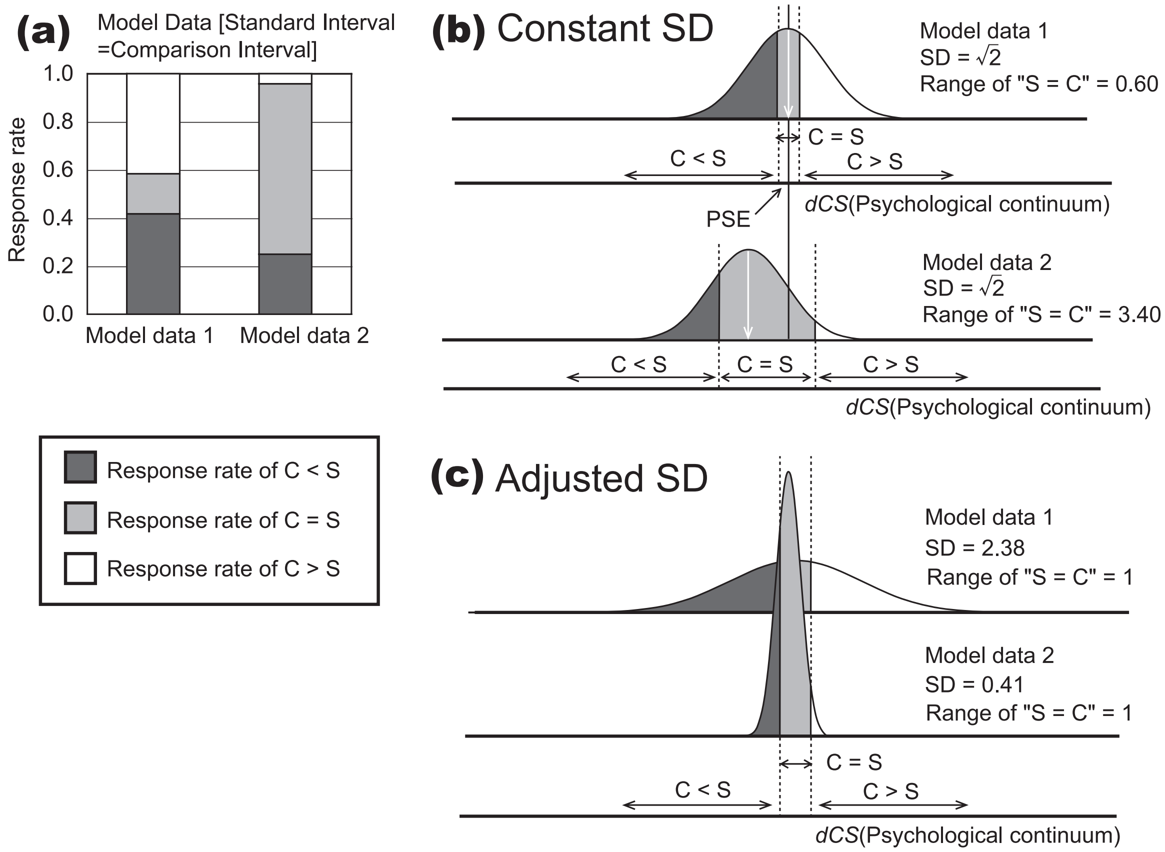

Fig. 7 Model data from a listener where comparison interval was physically equal in standard interval. Panel (a) shows response rate for each response category, which was the same as in actual Experiment II. Panel (b) illustrates $\overline{\psi_{dCS}}$ and $\sigma_{\psi_{dCS}}$ of model data 1 and 2 assuming of Thurstone's Case V. Panel (c) shows the case of assuming Thurstone's Case IV.

equation using Thurstone's law of comparative judgment [20]:

$$\sigma_{\psi_{dCS}} = \sqrt{\sigma_{\psi_C}{}^2 + \sigma_{\psi_S}{}^2 - 2r_{\psi_C\psi_S}\sigma_{\psi_C}\sigma_{\psi_S}}. \quad (3)$$

Let us assume the case of applying the law of comparative judgment for the two alternative forced choice (2AFC) test to data from three alternative forced choices that allow "equal" response to be added to the general 2AFC test. Bar graphs in Fig. 7(a) show model data 1 and 2 from repeated observations of a listener's responses when C and S have the same physical value. It can be assumed that the listener judges "C to be equal to S" when the dCS value calculated by Eq. (2) is within a certain range. If dCS is smaller than a certain "equal" range, C is judged to be smaller than S, and if dCS is greater than a certain "equal" range, C is judged to be greater than S. Applying Thurstone's case V to this process of judgment, a fixed value is obtained for $\sigma_{\psi_{dCS}}$ ($\sigma_{\psi_{dCS}} = \sqrt{1^2 + 1^2}$), since Thurstone's case V assumes that the coefficient of correlation r between ψ_C and ψ_S will be 0, and the standard deviation of ψ_C (σ_{ψ_C}) will be 1, which is equal to the standard deviation of ψ_S (σ_{ψ_S}) [20].

If $\sigma_{\psi_{dCS}}$ is assumed to be fixed at a certain value, to fit the observed data, one must change the range of the judgment that "C is equal to S" on a psychological continuum (see Fig. 7(b)). To change this range for "equal" judgment, a priori information that affects the criteria of judgment should be provided for each judgment. No such information was given to the listeners in Experiment II. Therefore, it is inappropriate to assume Thurstone's case V in the current case.

Without any a priori information, it is natural to assume that boundary points between the three judgment categories on a psychological continuum are fixed. To fit the dispersion model to the observed data under this assumption, the dispersion width of distributions for C and S should be changed. Applying Thurstone's Case IV, the value range for the "$C = S$" response category is fitted to ψ_{dSC}. This fitting modifies $\sigma_{\psi_{dCS}}$. In this application, $\sigma_{\psi_{dCS}}$ plays the role of an index of the psychological representation of the difference between stimuli. For example, if $\sigma_{\psi_{dCS}}$ is small, the value of the difference between stimuli falls into a narrow range on a psychological continuum. In addition, according to the fitting of the value range of "equal" judgments, the mean of ψ_{dCS} ($\overline{\psi_{dCS}}$) can be

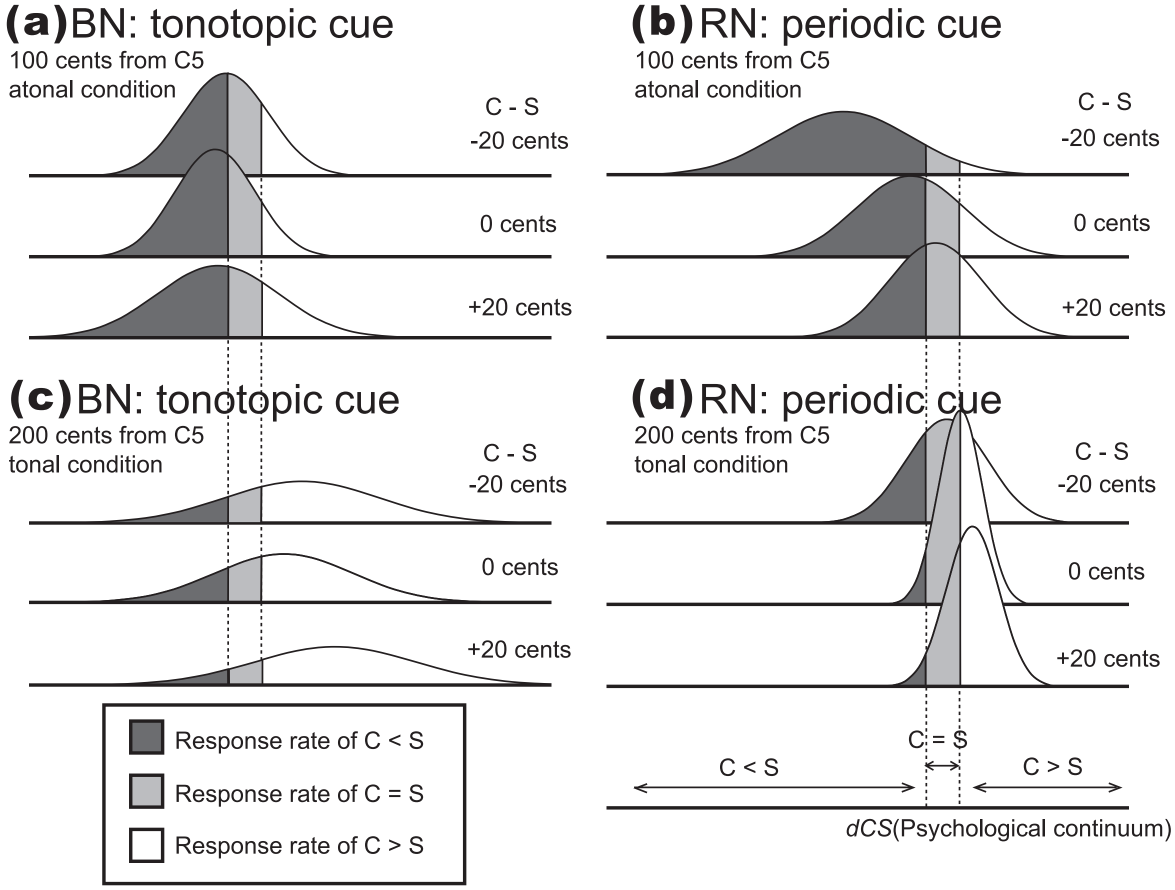

Fig. 8 $\overline{\psi_{dCS}}$ and $\sigma_{\psi_{dCS}}$ assuming Thurstone's Case IV of corresponding data shown in Fig. 6. Panel correspond to the respective four panels in Fig. 6.

calculated as a relative value throughout all comparisons. This mean value can be regarded as the psychological distance between ψ_{dCS} in each comparison. We use this model and review the difference in the psychological representation for each discriminal process.

3.4. Model Fit and Discussion

$\overline{\psi_{dCS}}$ and $\sigma_{\psi_{dCS}}$ for the standard intervals of 100 cents and 200 cents starting at C5 in Fig. 6 are illustrated in Fig. 8. The dark grey area corresponds to the response rates for "$C < S$," the light grey area corresponds to the response rates for "$C = S$," and the white area corresponds to the response rates for the "$C > S$" judgment.

An ANOVA was performed for $\sigma_{\psi_{dCS}}$ under all conditions to determine the effects of all factors (type of stimulus, in tonal or out of tonal, and deviation of comparison interval from standard interval) and their interactions. The effects of stimuli type and the interaction between type of stimuli and tonal framework were significant [$F(2, 24) = 28.14$, $p < 0.0001$. $F(1, 24) = 16.5843$, $p = 0.0004$, respectively]. It could be concluded that $\sigma_{\psi_{dCS}}$ in the RN condition was smaller than that in the BN condition, especially in the tonal framework.

It should be noted that, on comparing Fig. 8(a) to (b), the order of $\overline{\psi_{dCS}}$ does not systematically correspond to the increment of the physical value of the difference between the C and S stimuli in BN. The estimation of $\sigma_{\psi_{dCS}}$ for a stimulus is affected by the manner of responses to contiguous stimuli, in addition to the accuracy of discrimination. For this reason, it was inappropriate directly compare $\sigma_{\psi_{dCS}}$ in (a) with $\sigma_{\psi_{dCS}}$ in (b), or that in (c) with that in (d).

Nevertheless, it can be pointed out that the dispersion of the discriminal model in Fig. 8(d) appeared steeper than those under the other conditions. This phenomenon reflects that the response rate of "$C = S$" for RN in the tonal condition was increased compared to the other conditions in Fig. 6.

The tonal context was defined as the effect that made the representation of the pitch interval more precise in the previous study [19]. Considering the current results, tonal context depends on the periodic cue provided by each note. In the previous study, tones with six successive harmonic components were used in the experiments [19]. Both tonotopic and periodic information were available for judgment. $\sigma_{\psi_{dCS}}$ was decreased when the interval was under the tonal context. It can be assumed that the result of the previous study was due to the periodic information of the harmonic complexes.

4. GENERAL DISCUSSION

The results of Experiment I reconfirmed the importance of melodic contour, as demonstrated by the fact that the correct response rate for the CP lure condition was lower than that for the CD-PP lure condition in the 2AFC paradigm. When the global contour was shared between the two alternatives, they tended to be more confusable than in the case of unshared contours. It is worth noting that this melodic contour concept is slightly different from what is used by Dowling and Fujitani [1]. Their melodic contour meant a sequence of symbolic codes, i.e., a series of up-down directions denoted by plus/minus signs. In contrast, the current concept of the melodic contour includes melodic interval information, that is, the distances between adjacent notes. The redefinition of melodic contour can be presumed on the basis of the fact that the correct response rate under the CP-PP lure condition was significantly higher than that under the CP lure condition, although these two lure conditions involve the "identical" melodic contours if one follows Dowling and Fujitani's definition [1].

Moreover, the result did not support the main hypothesis of Experiment I that the property of each note exclusively corresponds to each melodic property, i.e., the tonotopic cue conveyed melodic contour information, while the periodic cue conveyed melodic interval information. It had been hypothesized that the periodic cue was inadequate to extract a melodic contour because of its circular nature. There was the possibility that the circularity of the periodic cue might bring about some confusion in defining the direction of the trajectory. However, the results indicated that the periodic cue sufficiently conveyed melodic contour information. This result might be affected by a few factors: the melodies used as stimuli had no large interval like the intervals used in the case of the tritone paradox [21]; the paradigm of Experiment I was to identify or recognize not the transposed melody, but 2AFC between the exactly transposed melody against the lure melody.

Furthermore, the result also revealed that the tonotopic cue could not convey sufficient information to discriminate the deviation of a semitone. In general musical scenes, the deviation of a semitone is distinctive. For listeners with a certain level of musical training, the semitone deviation often causes a feeling of something being "amiss," particularly in tonal music. The fact that the deviation of a semitone is not distinctive with only the tonotopic cue implies that the periodic cue could contribute to the recognition of melodic intervals in the tonal context more than the tonotopic cue.

In Experiment II, we investigated the functional difference between the tonotopic and periodic cues where the listeners were asked to discriminate the size relation between two melodic intervals in the tonal context given by a preceding diatonic scale. Consequently, the periodic cue helped to represent the melodic intervals as a steeper distribution on a psychological continuum, resulting in a higher probability of the listeners having a sense of equality between the two intervals when the intervals were reasonably identical in a musical sense, particularly in the tonal framework.

The results of the current study reverified the importance of melodic contour in the recognition of melodies, as claimed by Dowling and Fujitani [1]. However, the results of Experiments I and II suggested that a different aspect of the melodic contour should be considered in addition to the simple sequence of symbolic directional change described by Dowling and Fujitani [1]. The term "melodic contour" could be misleading if it imposed a bias to infer a rough transitional pattern on a frequency axis. It is periodic information, rather than tonotopic information, that serves the essential function in the recognition of melodies, particularly in a tonal framework. Dowling and Fujitani [1] also claim the importance of the tonal framework in the recognition of melodies on the basis of their experimental results. They reported that the rate of false recognition for the "Tonal Lure," that is, a shifted pattern in the same key as the original key, was prominent in spite of a semitone deviation in the relation between intervals.

The current study revealed that cues provided by each sound composing a melody played a significant role in the recognition of melodies.

5. SUMMARY

To investigate the cue in the recognition of a transposed melody, an experiment was carried out where the task was to recognize a transposed atonal melody using stimuli providing exclusively the tonotopic or periodic cue. The result revealed the importance of melodic contour, which was not simply symbolic up-down patterns but also reflected interval sizes. Although the tonotopic cue provided less accurate information than the periodic cue, both tonotopic and periodic cues could convey contour information. In Experiment II, the functional difference between the tonotopic and periodic cues was investigated in a tonal context and found to affect the perceptual judgment of the melodic interval sizes. A sharper perceptual representation was estimated for the stimuli providing the periodic cue than those providing the tonotopic cue. This observation suggested that the periodic cue provided a more robust tonal context than the tonotopic cue.

ACKNOWLEDGEMENT

This study was supported by Grant No. 17530529(C) from JSPS.

REFERENCES

[1] W. J. Dowling and D. S. Fujitani, "Contour, interval, and pitch recognition in memory for melodies," *J. Acoust. Soc. Am.*, **49**, 524–531 (1971).

[2] J. C. R. Licklider, "A duplex theory of pitch perception," *Experientia*, **7**, 128–133 (1951).

[3] D. Deutsch, "The tritone paradox: Effects of spectral variables," *Percept. Psychophys.*, **41**, 563–575 (1987).

[4] W. A. Yost, R. D. Patterson and S. Sheft, "A time domain description for the pitch strength of iterated ripple noise," *J. Acoust. Soc. Am.*, **99**, 1066–1078 (1996).

[5] W. A. Yost, "Pitch of iterated rippled noise," *J. Acoust. Soc. Am.*, **100**, 511–518 (1996).

[6] W. A. Yost, "Pitch strength of iterated rippled noise," *J. Acoust. Soc. Am.*, **100**, 3329–3335 (1996).

[7] B. C. J. Moore, "Frequency difference limens for short-duration tones," *J. Acoust. Soc. Am.*, **54**, 610–619 (1973).

[8] W. J. Dowling, "Scale and contour: Two components of a theory of memory for melodies," *Psychol. Rev.*, **85**, 341–354 (1978).

[9] W. J. Dowling and J. C. Bartlett, "The importance of interval information after long and short delays," *Psychomusicology*, **1**, 30–49 (1981).

[10] L. A. DeWitt and R. G. Crowder, "Recognition of novel melodies after brief delays," *Music Percept.*, **3**, 259–274 (1986).

[11] W. J. Dowling, "Tonal strength and melody recognition after long and short delays," *Percept. Psychophys.*, **50**, 305–313 (1991).

[12] W. J. Dowling, S. Kwak and M. W. Andrews, "The time course of recognition of novel melodies," *Percept. Psychophys.*, **57**, 136–149 (1995).

[13] F. Attneave and R. K. Olson, "Pitch as a medium: A new approach to psychophysical scaling," *Am. J. Psychol.*, **84**, 147–166 (1971).

[14] W. J. Dowling, "Context effects on melody recognition: Scale-step versus interval representation," *Music Percept.*, **3**, 281–296 (1986).

[15] W. L. Croonen, "Effects of length, tonal structure, and contour in the recognition of tone series," *Percept. Psychophys.*, **55**, 623–632 (1994).

[16] G. J. Balzano and B. W. Liesch, "The role of chroma and scalestep in the recognition of musical intervals in and out of context," *Psychomusicology*, **2**, 3–31 (1982).

[17] J. Wapnick, G. Bourassa and J. Sampson, "The perception of tonal interval in isolation and in melodic context," *Psychomusicology*, **2**, 21–37 (1982).

[18] A. Rakowski, "Intonation variants of musical intervals in isolation and in musical context," *Psychol. Music*, **18**, 60–72 (1990).

[19] M. Tsuzaki, "Effect of the preceding scale on melodic interval judgment in terms of equality and size," *Music Percept.*, **9**, 47–70 (1991).

[20] L. L. Thurstone, "A law of comparative judgment," *Psychol. Rev.*, **34**, 273–286 (1927).

[21] D. Deutsch, "Octave generalization of specific interference effects in memory for tonal pitch," *Percept. Psychophys.*, **13**, 271–275 (1973).

Toshie Matsui received Bachelor of Music and Master of Music from Kyoto City University of Arts in 2000 and 2003, respectively. She is a PhD candidate at Kyoto City University of Arts and a research member of JST CrestMuse project at Kwansei Gakuin University. She joins the Acoustical Society of Japan, the Japanese Society of Music Perception and Cognition, and the Japanese Psychonomic Society.

Minoru Tsuzaki After getting M. A. in psychology at Tokyo University in 1982, worked as a research assistant for Niigata University and Tokyo University for three years, respectively. In 1988, joined the first phase of ATR project (AVP) at Kyoto, and also worked for the second phase (HIP) and for the third phase (SLT). In 2004, moved to Kyoto City University of Arts as an associative professor. His research interest lies in time perception, timbre perception, and event perception with links to the basic auditory mechanism. Besides the Acoustical Society of Japan, joins the Acoustical Society of America, the Japanese Psychological Association, the Japanese Society of Music Perception and Cognition.