

Dynamic Cues in Key Perception

Rie Matsunaga (Corresponding author)

Department of Psychology, Hokkaido University
Kita 10 Nishi 7, Kita-ku, Sapporo, 060-0810, Japan
Tel: 81-11-706-3912 E-mail: matsunaga.rie@let.hokudai.ac.jp

Jun-ichi Abe

Department of Psychology, Hokkaido University
Kita 10 Nishi 7, Kita-ku, Sapporo, 060-0810, Japan
E-mail: abe@let.hokudai.ac.jp

Received: December 20, 2011

Accepted: February 1, 2012

Published: March 1, 2012

doi:10.5539/ijps.v4n1p3

URL: <http://dx.doi.org/10.5539/ijps.v4n1p3>

The research was supported by Grant-in-Aid for JSPS Fellows (19-11927).

Abstract

The traditional idea of pitch full-set alone cannot explain different key perceptions for melodies that consist of the same pitch full-set but differ in pitch sequence. Three experiments, in which presentation styles, participant groups, and stimulus sets were manipulated independently, traced the process of key development back from a final stage of key identification to earlier stages of listening to a melody. In all the experiments, two results were confirmed: First, key responses in earlier stages influenced those in later stages to the extent that subsequent tones correspond to scale tones of a previously interpreted key, revealing a phenomenon termed perceptual inertia. Second, when multiple keys were possible, listeners tended to perceive the diatonic key that can contain more pitch classes within a pitch set given at the point of time as stable scale tones of that key (i.e., tonic triad).

Keywords: Cognition, Music perception, Tonality, Musical training

1. Introduction

When a sequence of tones is organized perceptually as a “melody”, listeners perceive a key associated with the sequence, regardless of their ability to consciously name it. What kinds of properties in an arbitrary melody function as cues for perceiving the musical key? To date, a large number of studies addressing this issue, at least implicitly, rely upon the concept of “pitch full-set” of the whole melody (e.g., Abe & Hoshino, 1990; Balzano, 1982; Longuet-Higgins, 1987; Krumhansl, 1990). This idea holds that the set of all pitch classes present within a complete melody, regardless of the temporal ordering of these pitch classes, is critical to identifying the melody’s key. Considerable evidence supports this position (e.g., Abe & Hoshino, 1990; Krumhansl & Kessler, 1982; Schmuckler & Tomovski, 2005). Specifically, these studies have demonstrated that listeners who are familiar with Western music identify the key of a melody by assimilating all of its constituent pitch classes of a given pitch full-set into the Western diatonic tonal schema. Consider, for example, a tone sequence consisting of six pitch classes [C, D, E, G, A, B] (Note 1). Because all six pitch classes are scale tones of each of C major, G major, E minor, and A minor (Note 2) listeners presumably should identify the melody as one of the four keys.

Krumhansl (1990) further developed the idea that a distribution of occurrence frequency of pitch classes within a given pitch full-set serves as the dominant cue for perceiving a key. This hypothesis posits that a listener is very sensitive to the relative frequency (or duration) of pitch classes that occur in the given pitch full-set. Krumhansl and Schumuckler proposed a key-finding computational model based on the idea of pitch class distributions (the K-S algorithm model). This algorithm calculates correlation coefficients between each of 24 key profiles (c.f. Krumhansl & Kessler, 1982) and a pattern of occurrence frequency of each pitch class employed in a melody; the key profile yielding the highest correlation determines the preferred key.

Although powerful, the idea of pitch full-set cannot account for all listeners' key responses. Consider the two melodies shown in Figure 1 (from Matsunaga & Abe, 2009). Both have the same pitch full-set, [C, D, E, G, A, B], but they differ in the sequential arrangement of these pitch classes. If pitch full-set alone serves as the critical cue to key, then listeners should judge both melodies as having the same key. Indeed, this is predicted by the K-S algorithm. Calculating correlations between each key profile and a pattern of occurrence frequency of pitch classes in each of the two melodies, the K-S algorithm gives G major as the most preferred key for both melodies. This algorithm evaluates the likelihood of each of four diatonic keys for the pitch full-set [C, D, E, G, A, B], assuming that the frequency distribution of these six pitch classes is uniform and that all classes have the same time values: G major ($r=.83$), E minor ($r=.73$), C major ($r=.67$), and A minor ($r=.66$) within the presented melodies. However, Matsunaga and Abe (2009) showed that listeners did not judge both melodies to be in the same key (e.g., G major). Instead, their listeners interpreted Melody 1 as being in C major and Melody 2 as being in G major. This phenomenon, together with other findings, clearly reveal that the pitch full-set idea falls short of explaining effects of the temporal ordering of pitch classes on key perception (e.g., Brown, 1988; Brown, Butler, & Jones, 1994; Butler, 1989; Deutsch, 1984; Matsunaga & Abe, 2005, 2009; West & Fryer, 1990). Nevertheless, the idea of pitch full-set is basically a valid explanation for key perception; it can be seen as a fundamental aspect of key perception. Nevertheless, questions remain regarding how this idea of a pitch full-set can be further developed to address issues related to temporal order.

Two approaches to modifications of the pitch full-set idea have been proposed. One approach is that although the pitch full-set functions as a primary influence on key perception, one or several highly salient (local) sequential features in a pitch sequence levy secondary influences. Several studies proposed features such as a critical pitch interval and certain pitch classes located at a critical sequential position. For example, Butler and Brown (e.g., Brown, 1988; Brown et al., 1994; Butler & Brown, 1994) argued that one important factor in key perception involves the presence of pitch intervals that are rare within a diatonic set (i.e., the augmented fourth or the diminished fifth). Nonetheless, the idea of rare intervals is equivalent to the idea of the pitch full-set (cf. Matsunaga & Abe, 2009 for a detailed discussion). Others have suggested the importance of different cues to key perception; these cues include the presence of the perfect fifth and either the major third or the minor third (e.g., Cuddy, 1991; Huovinen, 2002), the minor second in the final position (Boltz, 1989), the ascending fourth or the descending fifth in the opening position (e.g., Vos, 1999), the pitch class in the final position of a sequence (e.g., Creel & Newport, 2002), or the pitch class in the opening position in conjunction with a pitch class in the final position (e.g., Cuddy, Cohen & Mewhort, 1981). Recently, Matsunaga and Abe (2009) evaluated these proposals. Using melodies that shared the same pitch full-set, they simultaneously assessed the relative contributions to listeners' key distinctions due to a wide range of local sequential features, such as an order of two adjacent pitch classes (this feature distinguishes between C->D and D->C), an order of two adjacent pitch classes in a particular position (this feature distinguishes between C in the 1st position->D in the 2nd position and C in the 2nd position->D in the 3rd position), a pitch class in a particular position (this feature distinguishes between C in the 1st and the 2nd positions), and so on. Their results led to the conclusion that none of the local sequential features proposed by the previous studies contributed significantly to key perception. The findings cast doubt on the idea that salient local sequence properties, taken together with pitch full-set, provide a satisfactory explanation of key perception.

The second approach to modifying a pitch full-set idea is found in Toiviainen and Krumhansl (2003)'s tone transition model. In this model, contributions of all constituent pitches within a given tone sequence are assessed on a 'two-tone transition' basis (i.e., one tone following another, such that C4->D4 is a different input from D4->C4). Thus, the distribution of two-tone transitions functions as an input to the key-finding model. The authors compared predictions of the tone transition model with predictions derived from models based on pitch class distributions (e.g., the K-S model). Contrary to their expectations, results indicated that the tone transition model did not provide a significantly better fit to listeners' psychological data than the pitch class distribution model did.

At present, no approach fully accounts for temporal order effects on key perception. Accordingly, in the present research, we consider a different approach to the issue of temporal order and key perception. As noted above, previous results (Matsunaga & Abe, 2009) indicate that local sequential features do not contribute significantly to differentiations among a few keys. Therefore, the present approach does not make local sequential features conditional, such as the ordering of two adjacent pitches or the particular serial location of a pitch class. Instead, our approach examines the way in which each given set of pitch classes at each point in time, compared to a given pitch full-set, may elicit an evolving sense of key as a melody unfolds.

Consistent with this approach, people tend to make rapid key decisions (Krumhansl, 1990). To investigate the

mechanism underlying this process, we used short tone sequences comprising six different pitches. In this study, a set of six pitches is defined as a ‘pitch full-set’, while a set of pitches beginning with a melody’s initial pitch and ending with the pitch at serial positions 5, 4, 3, and 2 is defined as a ‘pitch subset’. This is suggested in the bottom portion of Figure 1. Viewed from this perspective, the two melodies of Figure 1 clearly share the same pitch full-set [C, D, E, G, A, B] that corresponds to the six-tone sequence. However, pitch subsets are different between the two melodies (see labeled Stages 5, 4, 3, and 2 in the lower portion of Figure 1). For instance, look at pitch sets in Stage 2. A pitch set in Stage 2 is [C, G] in Melody 1, and a pitch set in Stage 2 contains different pitch classes [B, D] in Melody 2. Such different pitch subsets between the two melodies can be found in Stage 3, Stage 4, and Stage 5, respectively. Thus, these melodies consist of the same pitch full-set but differ in a transition of pitch subsets.

Over time, such transitions of the pitch subsets that ultimately comprise the pitch full-set may play a dominant role in key perception. Let us explain this idea more specifically. A pitch set at each point in time can momentarily govern key perception. Furthermore, key perception at various time points in the melody is not governed only by the pitch set at that time; it may also be influenced by the listener’s sense of key elicited at preceding stages in this process. Listeners can try to retain an already interpreted key as long as counter-evidence (e.g., a non-scale tone) against the current key interpretation does not occur. We refer to this phenomenon as perceptual inertia. In short, we hypothesize that perceptions of different keys that are elicited by different pitch (sub)sets at preceding stages in a sequence can, in turn, shape differentiations between a few keys evident at the final stage where the same pitch full-set occurs. To test the hypothesis, the present study focused on listeners’ evolving key perception. All experimental melodies consist of the same pitch full-set but differ in the order in which these pitches appear; in other words, these melodies differ from one another only in the transitions of pitch subsets. Thus, this study addressed the temporal order issue as a sequencing issue involving differences in transitions of pitch subsets.

In this study, three experiments were conducted. In each experiment, two types of analyses were performed. The first analysis examined whether the key percept evident in Stage 6 was already established prior to Stage 6. For instance, in light of Melody 1 and Melody 2 (Figure 1), we ask “How early in the listening process does a listener favor C major or G major in Melody 1 or in Melody 2, respectively?” To address such a question, we conditionalized the final key response to melodies (Stage 6); then we tested which keys participants judged in Stages 5, 4, 3, and 2; and tested whether there were tone sequences for which most participants successively retained the a final perception of key from the preceding stages prior to Stage 6. The second analysis attempted to isolate those aspects of pitch subsets (i.e., pitch set in each stage) that were most instrumental in differentiating among a few candidate keys. Reinterpreting the question in the light of Melody 1 and Melody 2 (Figure 1), this analysis asks what kind of differences in pitch subsets (between these melodies) lead to C major or G major perception, respectively. To address such a question, we enlisted discriminant analyses (cf. Hair, Anderson, Tatham, & Black, 1998; Klecka, 1980). A separate discriminant analysis was performed for each stage in order to examine the relationship between pitch subsets and participants’ key responses. Suppose, for example, that the results of discriminant analysis (at each stage) reveal that several pitch subsets contribute significantly to each of a few key response groups. In this case, it is possible to ascertain how participants arrive at these different keys (in the final stage where the same pitch full-set obtains) by tracing/identifying the distinctive pitch subsets shown to be influential prior to the final stage.

2. Experiment 1

Experiment 1 examined listeners’ key perception from the standpoint of a transition of pitch subsets within a pitch full-set. A set of 39 different melodic sequences with the same pitch full-set was constructed. Based on the results of our previous study (Matsunaga & Abe, 2009), we predict that when presented in the complete form, participants’ responses will tend to be distributed among C major, G major, and A minor. In Experiment 1, these sequences were presented to musicians with absolute pitch (AP), using a stage-dependent presentation style in which the number of tones within a presented segment increased with successive presentations.

Although AP musicians are often thought to be the special selected group, empirical studies have shown that tonal organization of AP musicians essentially shares the same characteristics with that of nonmusicians (e.g., Matsunaga & Abe, 2005; Temperly & Marvin, 2008). It is general that compared to responses of the nonmusician group, responses of the musician group can be less contaminated by extraneous factors. Moreover, in the musician group, AP possessors alone can offer the most direct key responses (i.e., key naming responses). Thus, we adopted AP musicians as participants for the present experiment.

2.1 Method

2.1.1 Participants

Participants were 15 musicians who possessed the ability of absolute pitch (13 women, 2 men; 18–22 years), 11 of whom were music majors. They had an average of 15.3 (range = 12–17) years of musical training. Eleven played the piano, and four played the electronic organ. The validity of AP participants was ensured by their self-report and their musical abilities (e.g., personal history of music lessons, performance ability, etc.).

2.1.2 Materials and Apparatus

Thirty-nine sequences of six pitches each [C, D, E, G, A, B] were used as stimuli. The pitch full-set was consistent with the following four Western diatonic keys: C major, G major, E minor, and A minor. All 39 tone sequences comprised the same full-set of six-pitch classes, but they differed in the temporal arrangement of these classes. Of the 39 tone sequences, 13 originated from pitch full-set I [C4, D4, E4, G4, A4, B4] and 26 came from II [D4, E4, G4, A4, B4, C5]. The only difference between pitch full-set I and II involved the register of C (C4 versus C5). We prepared two kinds of pitch full-sets to create stimulus melodies containing as many intervals as possible. Neither of these pitch full-sets included the intervals (± 6) (see Matsunaga & Abe (2009) for a detailed discussion on intervals (± 6)). Within the two pitch full-sets, there were 20 possible intervals between two pitches within the two pitch full-sets: (± 1), (± 2), (± 3), (± 4), (± 5), (± 7), (± 8), (± 9), (± 10), and (± 11) (Note 3). All tone sequences were monophonic isochronous melodies involving contiguous, non-overlapping, tones. They were created as MIDI files using sequencing software (Roland, “Cakewalk Pro Audio9”) installed on a Windows PC. All were presented at the same tempo, with identical tone durations (i.e., 0.6 s), leading to a total sequence duration of 3.6 s for each melody. The timbre of each pitch was that of an acoustic grand piano. Melodies were presented over two speakers in stereo (converted from mono MIDI files).

2.1.3 Procedure

Participants were tested individually. Seated in front of the two speakers, they were given a response sheet listing the 12 major and 12 minor key categories. Using a stage-dependent presentation paradigm, a given trial consisted of presenting a sequence in the following fashion: First, only the opening two pitches of the sequence (Stage 2) were presented; next, the first three pitches (Stage 3) were presented, and so on until the final presentation involved the whole sequence (Stage 6). Although the presentation style offered a natural way to trace a change of key perception as a melody unfolded, this experimental setting might encourage participants to maintain the key response that was given in a prior presentation within a trial. To guard against this possibility, participants were told the following: “You (i.e., the participant) must consider tone sequences given in each presentation as being independent when you select a key.” After three practice trials, 39 experimental trials were presented (one per melody) in randomized order for each participant.

2.2 Results and Discussion

Results are presented in two major sections. In the first section, we examined how early in the process participants favored a final key response in Stage 6. In a second section, we present findings based on discriminant analyses which assess the relative influence of different pitch subsets over stages on key responses.

2.2.1 Tracking Key Responses

Key categories with highly frequent responses were common to the two pitch full-sets I and II. Therefore, these data were pooled. Key responses in Stage 6 were predominantly limited to C major, G major, and A minor. Using a classification criterion requiring that seven or more of the 15 participants agreed on a key category for a tone sequence, the 39 sequences were classified into three groups, namely the C major group (16 tone sequences), the G major group (17 tone sequences), and a diverse response group (6 tone sequences).

After conditionalizing on final responses (i.e., C major and G major groups), we examined whether key responses for each group appeared in each of the stages preceding Stage 6. Distributions of key responses are shown in Figure 2 as a function of stage (Stages 6–2). In the C major group, C major responses constituted by far the largest proportion of all responses in all stages. In the G major group, G major responses constituted the largest proportion in all stages, and this proportion decreased as the stages traced back. These results indicated that the key categories with largest properties of responses in Stage 5, 4, 3, and 2 were consistent with those offered in Stage 6.

To determine whether key responses provided in the final stage were successively retained over earlier stages, we isolated tone sequences for which more than one-third of participants successively retained the final decision of key (i.e., C major for C major group, and G major for G major group) from stages prior to Stage 6. If

participants were strictly sensitive to changes of pitch set with the unfolding of a melody and they tended to shift key interpretations abruptly, there should be no evidence of key retention from Stages 5-2 to Stage 6. However, in the majority of tone sequences within the two key groups of interest, there was clear evidence of key retention. Specifically, in the C major group, seven tone sequences elicited the retention of the C major key response from Stage 5, four from Stage 4, zero from Stage 3, and four from Stage 2. In the G major group, six tone sequences elicited the retention of the G major response from Stage 5, eight from Stage 4, one from Stage 3, and two from Stage 2. Taken together, these results indicated that the final key perception (e.g., C major or G major) evident in Stage 6 resulted from participants retaining key interpretations made in preceding stages. In other words, AP musicians did not shift keys abruptly in response to changing pitch subsets; instead, they exhibited perceptual inertia in key perception. In other words, as a melody unfolds, they retain an early key interpretation as long as no counter-evidence (e.g., a non-scale tone) is encountered.

2.2.2 Discriminant Analysis

This section attempts to identify the critical pitch subsets that contribute to distinguishing among different key responses. To address this question, we used discriminant analyses. In this study, a separate discriminant analysis was performed for each of Stages 5, 4, 3, and 2 (Note 4). For performing each discriminant analysis, we focused on C major response, G major response, and A minor responses provided by all the participants for all the tone sequences. As a result, the sample sizes were 494, 472, 442, and 449 for Stage 5, 4, 3, and 2, respectively. The dependent variable of discriminant analyses in Stages 5-2 was groups corresponding key response categories of C major, G major and A minor. Independent variables in each discriminant analysis were each of all possible pitch sets (treated as categorical scales, i.e., 1 for presence and 0 for absence) for a given stage. In each discriminant analysis, after discriminant functions were developed, they were rotated orthogonally (just as in a factor analysis). Rotation facilitates substantive interpretation while preserving the original relations between structure coefficients of independent variables (e.g., Hair et al., 1998).

If results of the discriminant analysis for each stage indicate that a pitch subset satisfied two criteria, this pitch subset could be considered to be a determinant of key response differentiations in the corresponding stage. These criteria were (1) the pitch subset must be associated with one key response (e.g., C major) but not another (e.g., G major); and (2) the pitch subsets associated with each of the key responses should be transpositionally equivalent -- for different absolute pitches with the same tonal functions (e.g., doh (tonic), sol (dominant)).

Taken together, the results of discriminant analyses for Stages 4, 3, and 2 alone revealed that calculated discriminant functions were significant (Figures 3-5); however, discriminant analysis for Stage 5 did not reveal significant discriminant functions. Next, we describe the results of Stages 4, 3, and 2 in detail.

Before mentioning the results of each stage, it is important to first clarify how to interpret the results of discriminant analyses. In the left of each figure, the black cycles denote the "group centroid" of the three key response groups. Differences in the locations of centroids reflect dimensions along which the three key groups differ. For instance, the general technique is to start interpreting which groups can be distinguished by the most powerful significant function (i.e., function 1). In this regard, interpretations of discriminant functions are generally based on the directions (i.e., positive or negative) of the centroid value of each group.

In the right of each figure, each gray cycle indicates each independent variable with their respective structure coefficients. A structure coefficient is a simple correlation coefficient between each independent variable and each discriminant function, and it indicates the relative contribution of each independent variable. If an independent variable exhibits a structure coefficient of greater than $\pm .30$ and has the same sign in the discriminant function as that of the centroid of a certain group of a dependent variable, then this independent variable is said to positively contribute to this group (cf. Hair et al., 1998).

The results for discriminant analysis in Stage 4, in which pitch sets comprise four tones, appear in Figure 3. Based on the locations of group centroids, Function 1 can be interpreted as discriminating between C major (positive direction) and the other two keys (negative direction), whereas Function 2 differentiates G major (positive direction) from A minor (negative direction). Structure coefficients show that the positive direction of function 1 has two variables ([C, D, E, G] and [C, E, G, B]) exceeding $+ .30$. The positive direction of function 2 also has two different variables ([D, G, A, B] and [C, D, G, B]) exceeding $+ .30$. Taken together, these results suggest that [C, D, E, G] and [C, E, G, B] are associated with C major, while [D, G, A, B] and [C, D, G, B] are associated with G major. The [C, D, E, G] and [C, E, G, B] are interpretable as [doh, re, mi, sol] and [do, mi, sol, ti] in C major, respectively. The [D, G, A, B] and [C, D, G, B] are interpretable as [sol, doh, re, mi] and [fa, sol, doh, mi] in G major. The [C, D, E, G] of C major and the [D, G, A, B] of G major indicated that the four-pitch set for each key differed in pitch classes, but they nonetheless shared comparable tonal functions. These pitch sets

include a set of intervals $[(\pm 2), (\pm 3), (\pm 4), (\pm 5), (\pm 7)]$, $[(\pm 2), (\pm 3), (\pm 5), (\pm 8), (\pm 10)]$, or $[(\pm 2), (\pm 4), (\pm 5), (\pm 7), (\pm 9)]$ because there were two kinds of C (i.e., C4 and C5) in this study. These results suggest that the four-pitch sets with these three sets of intervals function as determinants of key differentiation in Stage 4. Again, this analysis also reveals the participants' sensitivity to A minor (Note 5); A minor is associated with [C, E, A, B], which is interpretable as [doh, mi, la, ti] in A minor.

The results for discriminant analysis in Stage 3 appear in Figure 4; these involve pitch sets that include the first three tones. Function 1 was interpreted as discriminating between C major (positive direction) and the other two keys (negative direction), and Function 2 was interpreted as discriminating between A minor (positive direction) and G major (negative direction). Structure coefficients show that the positive direction of function 1 has [C, E, G] and [D, E, G], and the negative direction of function 2 has [D, G, B] (its structure coefficients was slightly lower than -.30). The analysis showed that [C, E, G] and [D, E, G] are associated with C major, while [D, G, B] is associated with G major. The pitch sets of [C, E, G] and [D, E, G] are interpretable as [doh, mi, sol] and [re, mi, sol] in C major, respectively. [D, G, B] is interpretable as [sol, doh, mi] in G major, respectively. The [C, E, G] of C major and [D, G, B] of G major differed in pitch classes but shared tonal functions. The [C, E, G] and [D, G, B] include intervals $[(\pm 4), (\pm 3), (\pm 7)]$, $[(\pm 8), (\pm 3), (\pm 5)]$ or $[(\pm 9), (\pm 4), (\pm 5)]$. This discriminant analysis also reveals the participants' sensitivity to key information associated with A minor. A minor was associated with [C, E, B], [C, E, A] and [D, E, A], which is interpretable as [doh, mi, ti], [doh, mi, la] and [re, mi, la], respectively, in A minor.

Finally, the results of discriminant analysis for Stage 2 appear in Figure 5; these involve pitch sets that include the first two tones. Function 1 was interpreted as discriminating between A minor (positive direction) and the other two keys (negative direction), and Function 2 was interpreted as discriminating between C major (positive direction) and G major (negative direction). Structure coefficients show that the positive direction of function 2 has [C, E], [E, G], and [C, G], the negative direction of function 2 has also [D, B], [G, B], and [D, G]. Taken together, these results suggest that [C, E], [E, G], and [C, G] are associated with C major, while [D, B], [G, B], and [D, G] are associated with G major. The pitch sets of [C, E], [E, G], [C, G] are interpretable as [doh, mi], [mi, sol], and [doh, sol] in C major, respectively. On the other hand, [D, B], [G, B], and [D, G] are interpreted as [sol, mi], [doh, mi], and [sol, doh] in G major, respectively. These relationships between the two-pitch sets of C major and those of G major differ in pitch classes but share tonal functions. The [C, E] and [G, B] sets include pitch intervals (in semitone steps) of (± 4) or (± 8) ; [E, G] and [D, B] include (± 3) or (± 9) ; [C, G] and [D, G] include (± 7) or (± 5) . This discriminant analysis also reveals the participants' sensitivity to key information associated with a minor key, A minor. A minor was associated with [E, A], which is interpreted as [mi, la] in A minor.

A summary of the results in Stages 4, 3, and 2 revealed that several pitch subsets made reliable contributions to each of the three key responses. Across the three stages, these pitch subsets shared (± 7) and its inversion (± 5) , (± 4) and its inversion (± 8) , and (± 3) and its inversion (± 9) . These pitch intervals correspond to intervals that consist of the "tonic triad" of a diatonic scale; this suggests that, relative to other scale tones, the scale tones comprising a tonic triad (i.e., doh, mi, sol in a major mode; la, doh, mi in a minor mode) are more stable and exert greater influence on participants' interpretations than do other scale tones.

The present experiment worked backwards from the final decision of the AP musicians (Stage 6) between C major and G major by conditionalizing this decision (where the pitch full-set was shared). The results indicate that if an additional pitch in an unfolding melody can be interpreted as a scale tone of an already-interpreted key, then AP musicians prefer to remain in the current key interpretation, even when pitch subsets change. This confirms the idea that AP musicians exhibit perceptual inertia in perceiving a key. Moreover, the results of discriminant analyses pinpoint several pitch subsets as significantly contributing to each key within each corresponding stage. Common features distinguished these significant pitch subsets, suggesting that if several theoretically possible diatonic keys exist for a given sequence, then AP musicians prefer to perceive the sequence as being in a key that contains more scale pitches than the tonic triad (i.e., one that has greater stability). This confirms that AP musicians familiar with Western music are likely to have a diatonic tonal schema in which scale tones are differentiated in terms of perceptual stability. In conclusion, the present experiment suggests that the key perception of AP musicians is governed by the pitch set at each point in time that a new input tone occurs and is influenced by the key percepts elicited at preceding stages.

3. Experiment 2

The purpose of Experiment 2 was to establish the generality of the findings reported in Experiment 1. The procedure was similar to that of Experiment 1; however, the presentation style and the pitch full-set employed were different in Experiment 2.

It can be argued that the stage-dependent presentation style used in Experiment 1 encourages participants to retain already selected keys. That is, perhaps the evidence for perceptual inertia in AP musicians is simply an experimental artifact resulting from sequential presentations. To eliminate this possibility, Experiment 2 employed a presentation style that relied upon the complete set of stimuli. Stimulus sequences of all lengths were randomly presented in Experiment 2. If the results of Experiment 2 were similar to those of Experiment 1, it can be more firmly concluded that AP musicians generally manifest perceptual inertia.

Experiment 2 used a different pitch full-set than Experiment 1. This allows us to confirm whether the effect of temporal ordering of pitches on key perception is specific to the particular keys involved (C and G major) or is a more general phenomenon associated with various keys. The critical tone sequences in Experiment 1 were drawn from a pitch-full set [C, D, E, G, A, B] that relied on white-key notes. At least in Japanese culture, musicians who play a piano initially learned keys based on white-key notes and sometimes encountered black-key notes in pieces (cf. Miyazaki & Ogawa, 2006). The musicians may therefore have a white-key note bias. If so, then AP musicians should be more familiar with keys that contain a majority of scale tones corresponding to white-key piano notes, such as C major, G major, etc. Consequently, it is unknown whether AP musicians can distinguish between diatonic keys when a greater number of scale tones are consistent with black-key notes. To assess this, the pitch full-set [C#, D#, F, G#, A#, C] was used in Experiment 2. This pitch full-set is consistent with four diatonic keys: C# major, G# major, F minor, A# minor.

3.1 Method

Twelve AP musicians participated (11 women and 1 man; mean age = 24.8 years old). The musical abilities of each musician were sufficient to ensure that he/she possessed absolute pitch. They had an average of 18.6 years of musical training. All of them played the piano or the electric organ. None had participated in Experiment 1.

Thirty-nine sequences that consisted of six pitch classes [C#, D#, F, G#, A#, C] but differed in the temporal arrangement of these pitch classes were used as stimulus melodies in this experiment. This pitch full-set was created by shifting each constituent pitch in the pitch full-set [C, D, E, G, A, B] of Experiment 1 one semitone higher. Specifically, the pitch full-sets I [C#4, D#4, F4, G#4, A#4, C5] and II [D#4, F4, G#4, A#4, C5, C#5] were used. All pitches in each of the two pitch full-sets I and II can be interpreted as scale tones of the following keys: C# major, G# major, F minor, A# minor. Five sequences of different lengths were prepared from each of the 39 stimulus melodies. These sequences comprised the first two tones, first three tones, first four tones, first five tones, and all six tones from one stimulus melody. This resulted in a total of 195 sequences, of which 16 were redundant. Therefore, 179 unique sequences were used (26 sequences at a two-tone length, 38 at a three-tone length, 38 at a four-tone length, 38 at a five-tone length, and 39 at a six-tone length). The 179 sequences were presented in random order to participants. All other details of individual melody construction and presentation were the same as for Experiment 1; the procedure was essentially similar to that of Experiment 1.

3.2 Results and Discussion

3.2.1 Tracking Key Responses

Data were pooled across the two pitch full-sets. Key responses in Stage 6 were limited to C# major, G# major, and F minor. We identified 26 tone sequences for which more than six (of 12) participants agreed on key responses in Stage 6. Consensus key identifications emerged for 26 of the 39 six tone sequences: 22 tone sequences were judged as G# major, and four were categorized as C# major. The result showed that participants' key responses differentiated between G# major and C# major in Stage 6, confirming that the effect of temporal ordering of pitches on key perception arose in a pitch full-set of black-key notes as well as that of white-key notes.

Again, our analysis conditionalized participants' final responses (i.e., G# major and C# major groups) with the aim of examining the emergence of key responses in earlier stages (Stages 5-2) for the two key-final groups. Distributions of key responses for all stages are shown in Figure 6. In the G# major group, G# major responses were the most common across all stages; in the C# major group, C# major constituted the first or second largest proportion of responses across all stages.

To determine whether key responses provided in the final stage were successively retained from earlier stages, we identified tone sequences for which more than one-third of the participants successively retained the final key decision (i.e., C# major for the C# major group, G# major for the G# major group) from stages prior to Stage 6. In the 22 tone sequences of the G# major group, seven tone sequences elicited the retention of G# major from Stage 5, two from Stage 4, eight from Stage 3, and two from Stage 2. Similar results were seen in the C# major

group. Although the number of C# major tone sequences was small (four), one sequence elicited the retention of C# major from Stage 4. These results indicated that final key decisions (e.g., G# major or C# major) in Stage 6 resulted from retaining key decisions that were evident in earlier stages.

3.2.2 Discriminant Analysis

We sought to determine which pitch subsets led most participants to agreement on G# major or C# major, using a separate discriminant analysis for each of Stages 5, 4, 3, and 2. As in Experiment 1, here the C# major, G# major, and F minor categories were chosen as groups of dependent variables in the discriminant analyses in each of Stages 5, 4, 3, and 2.

Taken together, the results of the discriminant analyses for Stages 3 and 2 alone revealed that the calculated discriminant functions were significant; however, discriminant analyses for Stages 5 and 4 did not reveal significant discriminant functions. Next, we describe the results of Stages 3 and 2 in detail.

The results of discriminant analysis for Stage 3 appear in Figure 7. Based on the locations of the group centroids, Function 1 was interpreted as discriminating between G# major (negative direction) and the other two keys (positive direction), and Function 2 was interpreted as discriminating between F minor (positive direction) and C# major (negative direction). The structure coefficients suggested that the negative direction of Function 2 included [C#, D#, C], while the negative direction of Function 1 included [G#, A#, C], [D#, A#, C], and [F, G#, A#]. Taken together, C# major is associated with [C#, D#, C], while G# major is associated with [G#, A#, C], [D#, A#, C], and [F, G#, A#]. The pitch set of [C#, D#, C] corresponds to [doh, re, ti] in C# major, whereas [G#, A#, C], [D#, A#, C], and [F, G#, A#] correspond to [doh, re, mi], [sol, re, mi] and [la, doh, re] in G# major, respectively. Thus, significant pitch subsets of C# major and G# major do not share tonal functions. Nevertheless, it should be noted that both significant pitch sets [G#, A#, C] and [D#, A#, C] of G# major contain two constituent tones of a tonic triad of G# major. Finally, F minor was associated with [F, G#, C] and [D#, F, C], which are interpreted as [la, doh, mi] and [sol, la, mi] in F minor.

The results of discriminant analysis for Stage 2 appear in Figure 8. Function 1 was interpreted as discriminating between F minor (positive direction) and the other two keys (negative direction), and Function 2 was interpreted as discriminating between C# major (positive direction) and G# major (negative direction). The structure coefficients suggested that the positive direction of Function 2 included [C#, F] and [C#, G#], and the negative direction of Function 2 included [G#, A#], [D#, C], and [D#, G#]. Taken together, [C#, F] and [C#, G#] are associated with C# major, while [G#, A#], [D#, C], and [D#, G#] are associated with G# major. The pitch sets of [C#, F] and [C#, G#] are interpreted as [doh, mi] and [doh, sol] in C# major, respectively. [G#, A#], [D#, C], and [D#, G#] are interpreted as [doh, re], [sol, mi] and [sol, doh] in G# major, respectively. The [C#, G#] of C# major and [D#, G#] of G# major differ in pitch classes but share tonal functions. The [C#, G#] and [D#, G#] include (± 7) and (± 5), as there were two kinds of C# (i.e., C#4 and C#5) in these sequences. The results in the major keys were consistent with the results seen in F minor. F minor was associated with [F, C] and [F, G#], which is interpreted as [la, mi] and [la, doh] in F minor.

In summary, this experiment basically replicates the results of Experiment 1 and confirms its generalizability. In Experiment 2, the experimental setting alone is unlikely to encourage participants to carry over key interpretations from preceding stages within a stimulus melody. In spite of it, key responses provided in the preceding stages continued to influence key responses in later stages. It is clear that the perceptual inertia elicited in AP musicians here cannot be attributed to an experimental artifact. In other respects, the results of discriminant analyses pinpoint several pitch subsets as significantly contributing to each of few keys within each corresponding stage. The results of discriminant analyses are also consistent with those of Experiment 1: If there are theoretically possible diatonic keys for a given tone sequences, AP musicians prefer to perceive the sequence as being in a key that offers more scale tones (within a sequence) as members of tonic triad of a key. Finally, this experiment also establishes that AP musicians have the same perceptual tendency in distinguishing among candidate keys when the keys correspond to black-key notes (in a piano) as when they correspond to white keys.

4. Experiment 3

Experiment 3 examined non-musicians' key perception with the same presentation style (i.e., stage-independent presentation) and the same pitch full-set (i.e., [C#, D#, F, G#, A#, C]) as in Experiment 2. The aim of this experiment was to ascertain if results provided by non-musicians were equivalent to those provided by AP musicians (in Experiment 2).

To infer a non-musician's perceived key, we used the "final-tone extrapolation method" (Abe & Hoshino, 1990), which assumes that the tone selected as a final tone is likely to be the tonic or the nuclear tone of the perceived

key. In this experiment, we asked non-musicians to select the most plausible final tone from C#, G#, F, and A# categories. All the four tones were equivalent to tonics of diatonic keys for the pitch full-set that was used [C#, D#, F, G#, A#, C]. The main reason for this restriction of key categories was to ensure a satisfactory sample size for discriminant analysis. Although most listeners perceive a given tone sequence to be in one of the set of keys that contain all pitches of the sequences as scale tones, non-musicians typically are more variable than musicians because they tend to be more influenced by extraneous factors (Matsunaga & Abe, 2005). For instance, if non-musicians had to select a final-tone from the 12 tone categories within one octave, their responses might concentrate about the keys of C#, G#, F, and A#, but the degree of concentration would be relatively low and might not be sufficient to allow reliable analysis. To obviate this possibility, we chose to constrain participants' responses to four categories, as outlined.

4.1 Method

The participants were 20 undergraduate students (6 women and 14 men; mean age = 25.7 years old). None of the participants had formal music education. The materials and apparatus were identical to those used in Experiment 2. All participants were tested individually. They were given a MIDI keyboard (YAMAHA CBX-K1XG) in which four notes (C#, F, G#, A#) were marked. In a trial, participants heard a tone sequence twice. After the two presentations, participants were allowed to listen to a given tone sequence as many additional times as they requested. From among four tone categories (C#, F, G#, A#), participants were asked to orally respond with one tone as the most plausible final tone, by striking one of the four labeled notes in the MIDI keyboard. After three practice trials, 179 experimental trials were randomly presented for each participant.

4.2 Results and Discussion

4.2.1 Tracking Final-tone Responses

The data were pooled across the two pitch full-sets. In Stage 6, either G# or C# were selected as final tones for all stimulus melodies (i.e., with the same pitch full-set). There was a consensus among more than 10 participants for tone C# in 13 tone sequences and for tone G# in five tone sequences. We focused on the C# group and the G# group. Distributions of key responses for all stages for each group are shown in Figure 9. In the two groups, final tones (i.e., C#, G#) in Stage 6 were more common across all stages. To determine whether final-tone responses provided in the final stage were successively retained from earlier stages, we identified tone sequences for which more than one-third of the participants successively retained the final decision of their final-tone responses (i.e., C# for C# group, G# for G# group) from stages prior to Stage 6. In the 13 tone sequences of the C# group, eight tone sequences elicited a retention of the C# response from Stage 5, one from Stage 4, one from Stage 3, and two from Stage 2. In the five tone sequences of G# group, four tone sequences elicited a retention of the G# response from Stage 5, and one from Stage 4.

4.2.2 Discriminant Analysis

We sought to determine what critical content features of the pitch subsets led most participants to an agreement on the final tone of C# or G#, using separate discriminant analyses for each of Stages 5, 4, 3, and 2. Although it was possible to use all four key categories as groups of a dependent variable, non-musicians' responses were arguably more influenced by extraneous factors. Thus, because they constituted more than 25 percent of all responses (i.e., above chance-level), C# and G# responses were chosen as groups of a dependent variable. F and A# responses were removed. The numbers of observations for each stage satisfied sample size criteria for discriminant analysis.

The results of discriminant analysis for Stage 4 alone revealed that the calculated discriminant functions were significant; however, DA for Stages 5, 3, and 2 did not reveal significant discriminant functions. In the following passage, we describe the results of the Stage 4 analysis in detail.

The results of the discriminant analysis for Stage 4 appear in Table 1. Based on the values of the group centroids of the two keys and values of structure coefficients, the following interpretation is offered. The selected final tone C# is associated with [F, G#, A#, C] and [C#, D#, F, C], whereas the final tone G# is associated with [D#, G#, A#, C], [C#, D#, G#, C], and [C#, D#, A#, C]. Although this cannot be directly confirmed, both of these final tones (C#, G#) can be considered as tonics of major (not minor) keys according to the following rationale: C# major and G# major (not C# and G# minor) are equivalent to diatonic keys for the pitch full-set employed here. Therefore, the pitch sets [F, G#, A#, C] and [C#, D#, F, C] are interpreted as [mi, sol, la, ti], [doh, re, mi, ti] in C# major, respectively. Meanwhile, the pitch sets of [D#, G#, A#, C], [C#, D#, G#, C], and [C#, D#, A#, C] are interpreted as [sol, doh, re, mi], [fa, sol, doh, mi], and [fa, sol, re, mi] in G# major, respectively. Although the significant pitch subsets of C# major and those of G# major do not share tonal functions, it should be noted that

all the significant pitch sets of C# major and those of G# major contain two constituent tones of a tonic triad of C# major and G# major, respectively.

In summary, although the data were less clear-cut than those of AP musicians, non-musicians showed similar results. Again, using a stage-independent presentation style that was hence unlikely to encourage artificial key maintenance, Experiment 3 revealed that non-musicians nonetheless showed a tendency to retain the final tone selected in earlier stages when providing a final response in Stage 6. Moreover, the results of the discriminant analysis for Stage 4 suggest that when there were theoretically possible diatonic keys for a given tone sequences, non-musicians preferred to perceive the sequence as being in a key that contains more tones of the sequence as more stable scale tones. This confirms that, as with AP musicians, non-musicians possess a diatonic tonal schema in which scale tones differ in perceptual stability.

5. General Discussion

There is a general consensus that a pitch full-set functions as an important cue for key perception. However, the original idea of the pitch full-set does not entirely account for the finding that different key perceptions emerge for melodies that consist of the same six-pitch full-set but differ in the temporal arrangements of these pitches (e.g., Matsunaga & Abe, 2005). The present study addressed the temporal order issue by introducing the concept of changing pitch subsets (i.e., changing pitch set at each point) as an important property of melodic structure. Specifically, the three experiments tracked back to earlier listening stages to reveal the timeline of a key differentiation process with the changes of pitch (sub) sets.

In experiments with AP musicians, the first analysis revealed that these listeners differentiated among a few key categories even in stages prior to the final stage. Moreover, these skilled listeners preferred to retain a key that was interpreted earlier so that subsequent tones would be interpreted as scale tones within the retained key (Experiments 1 & 2). These results were not caused by experimental artifacts associated with sequential presentations. Instead, the results reflect AP musicians' mental characteristics (Experiment 2). Thus, we conclude that AP musicians exhibit perceptual inertia in perceiving a key. Moreover, the results of the discriminant analysis clearly indicated that, given multiple diatonic keys, AP musicians gravitated toward the key containing the most constituent tones within the given pitch set as stable scale tones (e.g., tonic triad tones), as shown in Experiment 1. Similar results emerged in diatonic key selections where most scale tones conformed to black-key notes (Experiment 2).

Results of Experiments 1 and 2 suggest that AP musicians possess a diatonic tonal schema, where scale tones differ in perceptual stability and tonal importance. The finding is consistent with the well-known tonal hierarchy theory (Krumhansl, 1990). However, it is worth noting that these experiments extend the validity of this theory using a methodology that differs from that employed in the majority of published research (i.e., the probe tone method).

According to common opinion, AP musicians represent a very small and select group and hear music in a fundamentally different way from the majority of listeners. Experiment 3 challenges this opinion. Experiment 3 examined the key perceptions of non-musicians. The results of this experiment concur with other researchers' results (e.g., Matsunaga & Abe, 2005; Temperly & Marvin, 2008) in indicating that characteristics of the tonal organization of non-musicians are basically equivalent to those of AP musicians. Specifically, like musicians, non-musicians exhibit perceptual inertia in key perception. Additionally, like musicians, non-musicians rely heavily on certain prominent scale tones with a hierarchically structured tonal schema. In spite of these commonalities, non-musicians differed from AP musicians in evidencing greater individual differences and more instability in their tonal schema. These differences are not surprising and are likely due to the fact that non-musicians are more influenced by extraneous factors than musicians (Matsunaga & Abe, 2005). Indeed, similar differences are often reported in other psychological experiments that compare the performance of experts with that of novices. In the present case, this does not negate our claim. In short, we infer that these skill differences are not due to essential characteristics of key perception.

To summarize our findings more concretely, we return to some sample Melodies 1 and 2 (cf. Figure 1) used in Experiment 1. Figure 10 summarizes how the current approach views changes in participants' key responses to each of the two melodies over different stages. For these melodies, performance in Stage 6 confirms that the majority of participants identified C major for Melody 1 and G major for Melody 2. Similar tendencies of key response are shown in the stages prior to Stage 6, suggesting that key perception in Stage 6 was partially shaped by key interpretations in Stages 5, 4, 3, and 2. This means that, in each stage, key interpretations were governed by the pitch set in the current time point and were influenced by persisting key interpretations from earlier stages. This is clearly illustrated in the responses at Stage 4. In this stage, judgments favoring C major for Melody 1

drop slightly, while A minor judgments suddenly appear at the same time. This may reflect the fact that the three-pitch set of Melody 1 changes at this point to include A4 [E4, G4, C5, A4], meaning that this four-pitch set included the most predominant scale tone in a hierarchical structure of A minor. Nonetheless, in this stage 10 participants continued to judge the sequence as conveying C major rather than A minor. The latter outcome could suggest a perceptual tendency toward maintaining a key endorsed in preceding stages. On the other hand, in Melody 2 the initial three-pitch set [D4, B4, C5] changed to [D4, G4, B4, C5] in Stage 4 by adding G4, namely the doh, tonic of G major. This means that Stage 4 for Melody 2 contains all tones of a tonic triad of G major, and consequently many more participants selected G major in this stage than in Stage 3.

To date, various key-finding models have been proposed (e.g., Bharucha, 1988; Krumhansl, 1990; Schmuckler & Tomovski, 2005; Temperly, 2001, 2008; Tillmann, Bigand, & Bharucha, 2000; Toiviainen & Krumhansl, 2003; Vos & Van Geenan, 1996; Yoshino & Abe, 2004). Some models are rule-based models, and others are connectionist models (i.e., neural network models). In the present study, we obtained empirical data that listeners exhibited perceptual inertia in key perception. Considering the perceptual inertia, it may be likely that the connectionist model is more potential approach over the rule-based models for explaining a human listener's key perception. This is because the connectionist model, such as the recurrent neural network (e.g., Elman, 1990), is able to incorporate the concept of dynamic temporal mechanism where outputs at earlier time points cumulatively influence those at later time points.

We note that the methodology of the present study does not focus directly upon local ordering of pitches within a pitch subset. This was not our primary aim. Instead, our purpose was to verify the idea that different key perceptions, led by a transition of pitch (sub)sets over preceding points, differentially influence later (i.e., final) key judgments of melodies composed of the same pitch full-set. In light of this, our stimulus melodies were designed to have the same pitch full-set. Thus, pitch subsets within stages prior to Stage 6 did not systematically consist of identical pitches across different melodies of the same length. Consequently, given this design, it is not possible to systematically analyze local temporal order effects within each pitch subset. For this reason, we treated pitch subsets as independent variables in the discriminant analyses. Accordingly, from the results of the discriminant analyses, we cannot ascertain if order differences obtain within a pitch subset. Furthermore, it should be noted that a given six-tone sequence holds a nesting structure of various pitch subsets. In other words, a six-pitch set comprises a five-pitch subset, a four-pitch subset, etc. A five-pitch set comprises a four-pitch subset, etc. When a pitch subset is made smaller, it finally comes down to the smallest unit of temporal order, i.e., a local sequential feature. Our previous study (Matsunaga & Abe, 2009) showed that none of the minimal units (i.e., local sequential features) significantly contributed to key response differentiations. This finding was the theoretical basis of the present study's aim and experimental design. Nonetheless, in light of the complexity of the temporal nesting of pitch sets, follow-up examinations that pinpoint various other factors that may contribute to perceptual inertia in key perception will be necessary. We think that such research is one of the challenges of the future.

In conclusion, the present findings suggest that, regardless of presentation style, participants' musical training, or stimulus sets, pitch (sub)set transitions play a dominant role in key perception. In other words, emerging key percepts, led by a transition of pitch (sub) sets over preceding positions, influence key judgments at a final stage where the same pitch full-set occurs.

References

- Abe, J., & Hoshino, E. (1990). Schema driven properties in melody cognition: Experiments on final tone extrapolation by music experts. *Psychomusicology*, 9, 161-172.
- Balzano, J. G. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, Mind, and Brain* (pp. 321-351). New York: Plenum Press.
- Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, 5, 1-30.
- Boltz, M. (1989). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 749-761. <http://dx.doi.org/10.1037/0096-1523.15.4.749>
- Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, 5, 219-250.
- Brown, H., Butler, D., & Jones, M. R. (1994). Musical and temporal influences on key discovery. *Music Perception*, 11, 371-407.

- Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, 6, 219-242.
- Butler, D., & Brown, H. (1994). Describing the mental representation of tonality in music. In R. Aiello (Ed.), *Musical perceptions* (pp.191-212). New York: Oxford University Press.
- Creel, S. C., & Newport, E. J. (2002). Tonal profiles of artificial scales: Implications for music learning. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition*. Sydney, Australia: Adelaide: Causal Productions.
- Cuddy, L. L. (1991). Melodic patterns and tonal structure: Converging evidence. *Psychomusicology*, 10, 104-126.
- Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 869-883. <http://dx.doi.org/10.1037/0096-1523.7.4.869>
- Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha, and Krumhansl. *Journal of Experimental Psychology: General*, 113, 413-416. <http://dx.doi.org/10.1037/0096-3445.113.3.413>
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211. http://dx.doi.org/10.1207/s15516709cog1402_1
- Hair, J. F., Anderson, E. R., Tatham, L. R., & Black, W. C (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Huovinen, E. (2002). *Pitch class constellations: Studies in the perception of tonal centrality*. Turku, Finland: Suomen Musiikkiteollinen Seura.
- Klecka, R. W. (1980). *Discriminant analysis. Quantitative Applications in the Social Sciences Series (No. 19)*. Thousand Oaks, CA: SAGE publications.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89, 334-368. <http://dx.doi.org/10.1037/0033-295X.89.4.334>
- Longuet-Higgins, H. C. (1987). *Mental processes: Studies in cognitive science*. London: MIT Press.
- Matsunaga, R., & Abe, J. (2005). Cues for key perception of a melody: Pitch set alone? *Music Perception*, 23, 153-164. <http://dx.doi.org/10.1525/mp.2005.23.2.153>
- Matsunaga, R., & Abe, J. (2009). Do local properties function as cues for musical key identification? *Japanese Psychological Research*, 51, 86-97. <http://dx.doi.org/10.1111/j.1468-5884.2009.00391.x>
- Miyazaki, K., & Ogawa, Y. (2006). Learning absolute pitch by children: A cross-sectional study. *Music Perception*, 24, 63-78. <http://dx.doi.org/10.1525/mp.2006.24.1.63>
- Schmuckler, M. A., & Tomovski, R. (2005). Perceptual tests of an algorithm for musical key-finding. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1124-1149. <http://dx.doi.org/10.1037/0096-1523.31.5.1124>
- Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.
- Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive Science*, 32, 418-444. <http://dx.doi.org/10.1080/03640210701864089>
- Temperley, D., & Marvin, W. E. (2008). Pitch-class distribution and the identification of key. *Music Perception*, 25, 193-212. <http://dx.doi.org/10.1525/mp.2008.25.3.193>
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107, 885-913. <http://dx.doi.org/10.1037/0033-295X.107.4.885>
- Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception*, 32, 741-766. <http://dx.doi.org/10.1068/p3312>
- Vos, P. G. (1999). Key implications of ascending fourth and descending fifth openings. *Psychology of Music*, 27, 4-18. <http://dx.doi.org/10.1177/0305735699271002>
- Vos, P. G., & Van Geenen, E. W. (1996). A parallel-processing key-finding model. *Music Perception*, 14, 185-224.

West, R., & Fryer, R. (1990). Rating of suitability of probe tones as tonics after random orderings of note of the diatonic scale. *Music Perception*, 7, 253-258.

Yoshino, I., & Abe, J. (2004). Cognitive modeling of key interpretation in melody perception. *Japanese Psychological Research*, 46, 283-297. <http://dx.doi.org/10.1111/j.1468-5584.2004.00261.x>

Notes

Note 1. Square brackets denote a set of pitches, regardless of the temporal ordering of the pitches. For example, the sequence C-D-E and the sequence E-D-C are represented as the same pitch full-set (i.e., [C, D, E]).

Note 2. In this paper, we did not consider harmonic minor and (ascending and descending) melodic minor scales.

Note 3. Intervals are denoted by positive integers for ascending intervals and by negative integers for descending intervals in semitones. For example, the ascending major third and the descending major third are denoted as (+4) and (-4), respectively.

Note 4. A separate discriminant analysis was not performed for Stage 6 because the six-tone pitch set of Stage 6 was equivalent to a pitch-full set that was shared among all the stimulus tone sequences.

Note 5. It was necessary to distinguish between the major and minor modes because the two major modes differ in the sequence of intervals (in semitones) between adjacent tones. Because of this, tonal functions of intervals in major keys are not always equivalent to those in minor keys even though the intervals of major keys and minor keys are the same.

Table 1. Experiment 3: Result of discriminant analysis for stage 4

Group centroids of discriminant functions		Structure coefficients	
Final tone type	Function 1	Independent variables	Function 1
Tone C#	-.150	C#D#FG#	-.215
Tone G#	.423	C#D#FA#	-.150
		C#D#FC	-.331
		C#D#G#C	.383
		C#D#A#C	.376
		C#FG#A#	-.016
		C#FG#C	-.213
		C#FA#C	.065
		D#FG#A#	.039
		D#FA#C	-.149
		D#G#A#C	.608
		FG#A#C	-.371

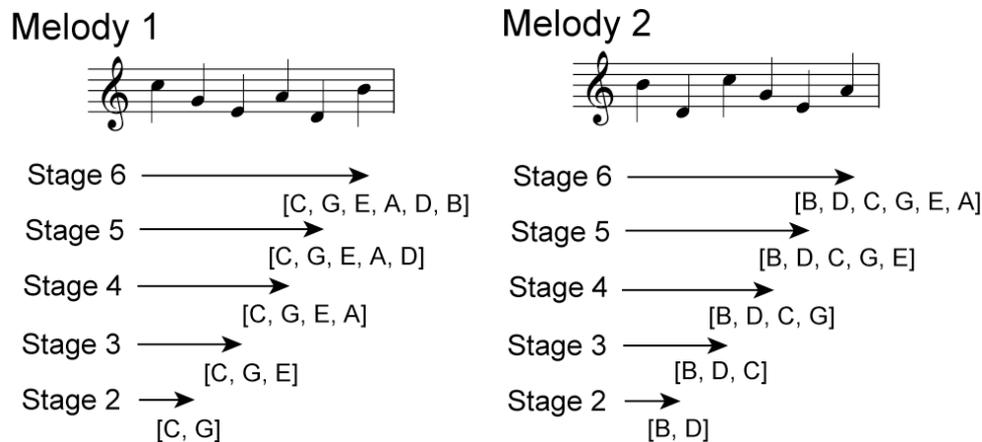


Figure 1. Melodies consisting of the same pitch full-set [C, D, E, G, A, B] but differing in pitch sequence. Arrows show pitch sets in each stage

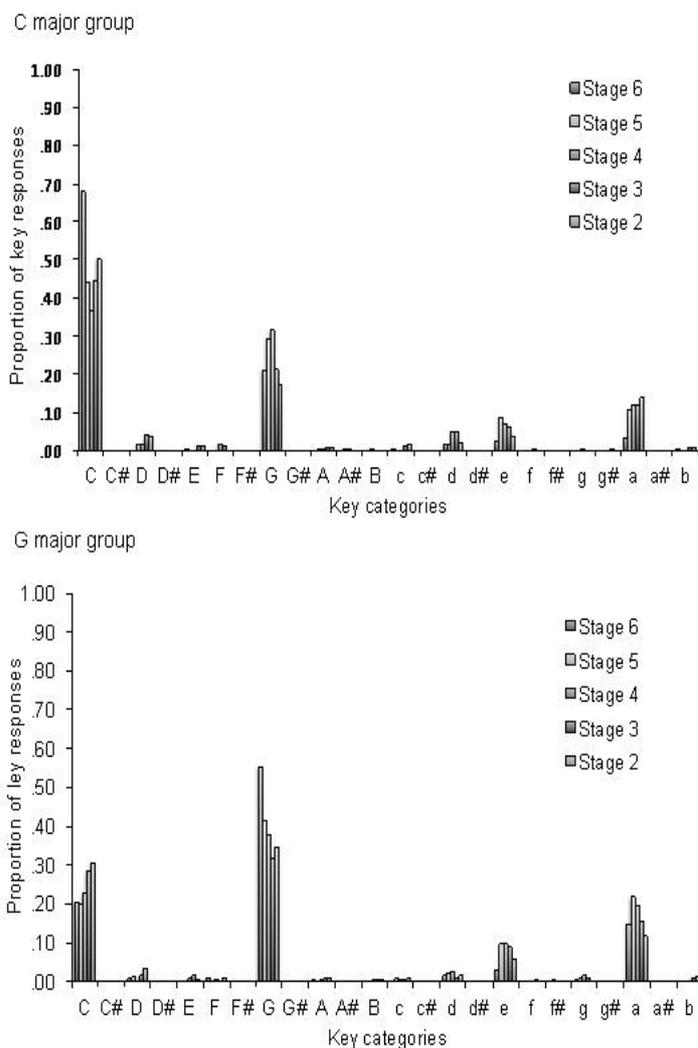


Figure 2. Distributions of key responses for each of the C major and the G major groups
 Capital letters and small letters denote major and minor keys, respectively

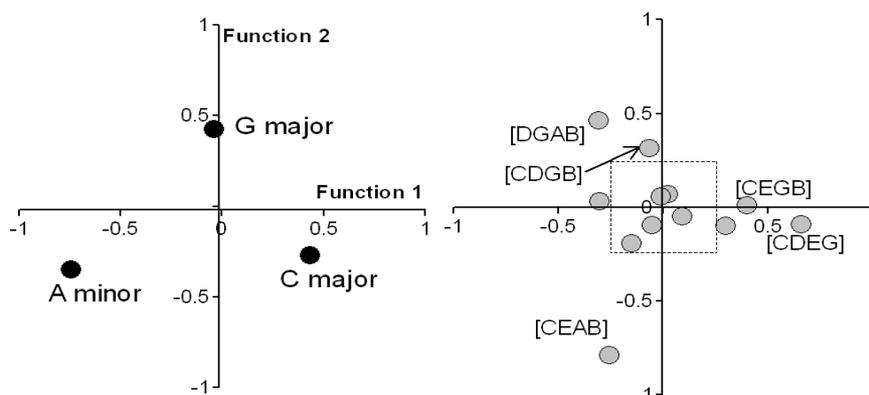


Figure 3. Experiment 1: Results of discriminant analyses for Stage 4

Black cycles indicate dependent variables (left), and gray cycles indicate independent variables (right). The dotted box indicates the boundary of $\pm .30$ of a structure coefficient.

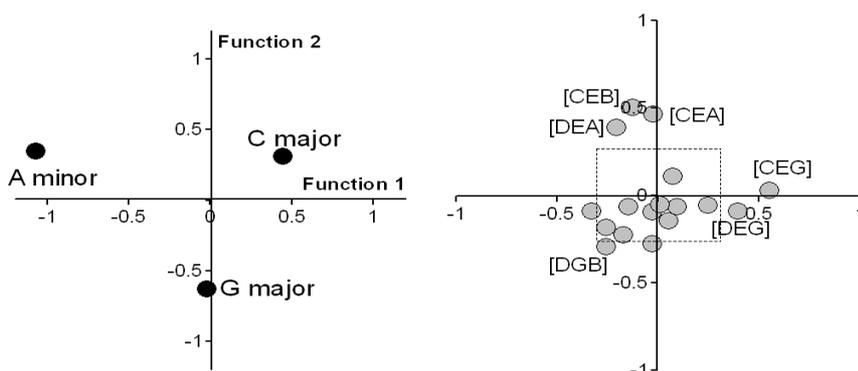


Figure 4. Experiment 1: Results of discriminant analyses for Stage 3

Black cycles indicate dependent variables (left), and gray cycles indicate independent variables (right). The dotted box indicates the boundary of $\pm .30$ of a structure coefficient.

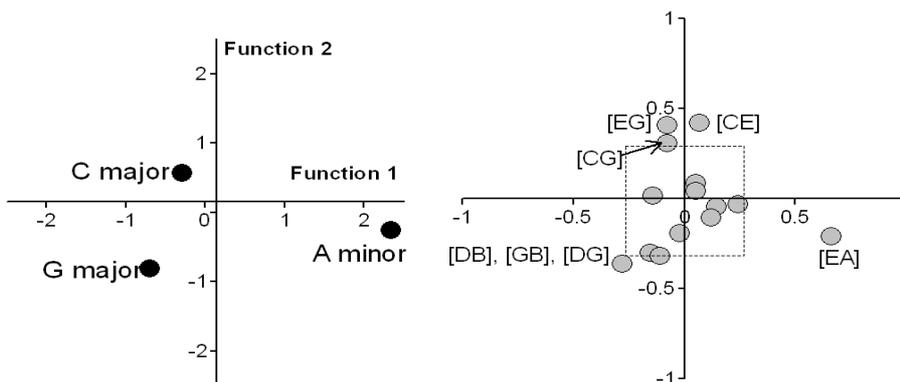


Figure 5. Experiment 1: Results of discriminant analyses for Stage 2

Black cycles indicate dependent variables (left), and gray cycles indicate independent variables (right). The dotted box indicates the boundary of $\pm .30$ of a structure coefficient.

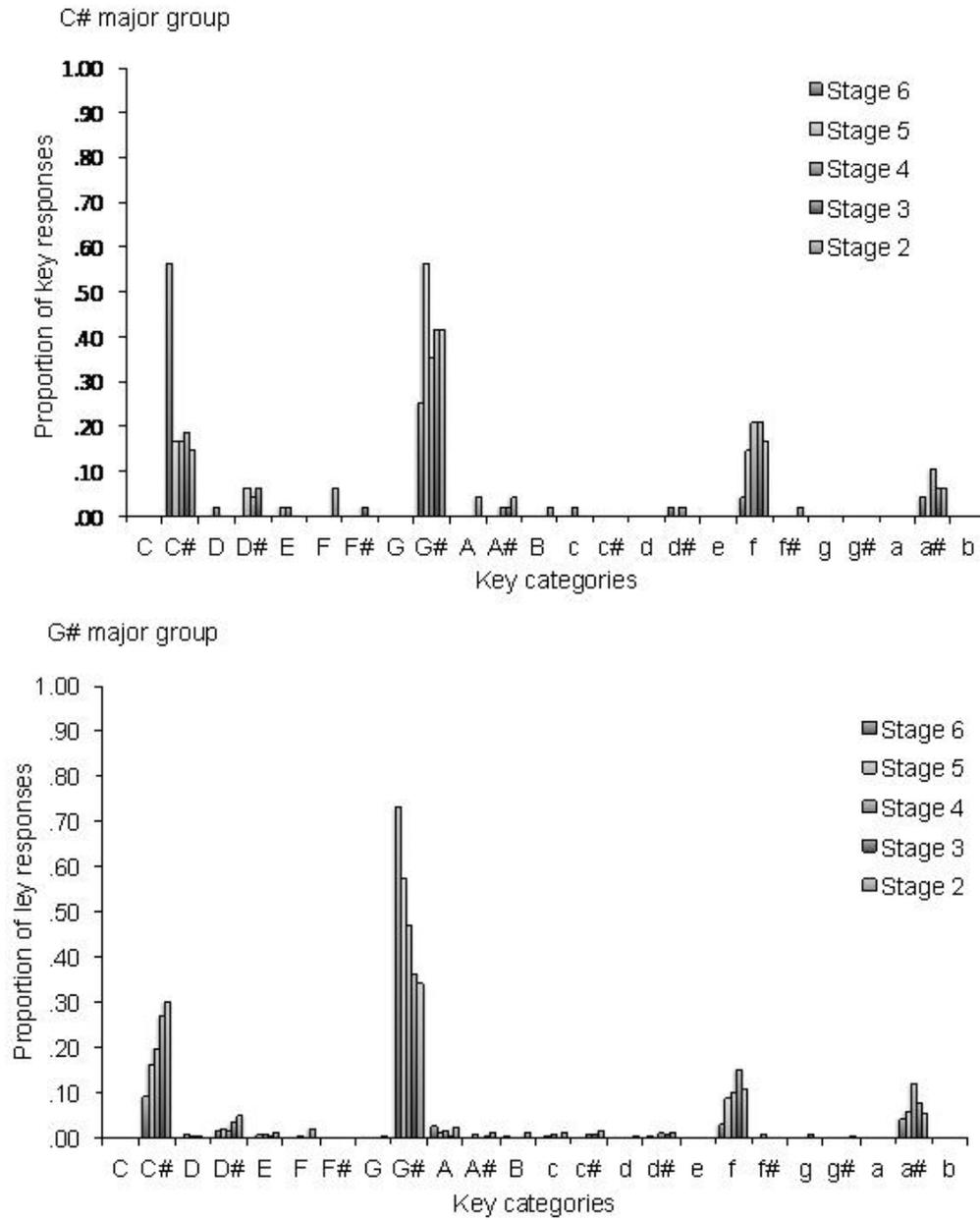


Figure 6. Distributions of key responses for each of the C# major and the G# major groups
Capital letters and small letters denote major and minor keys, respectively.

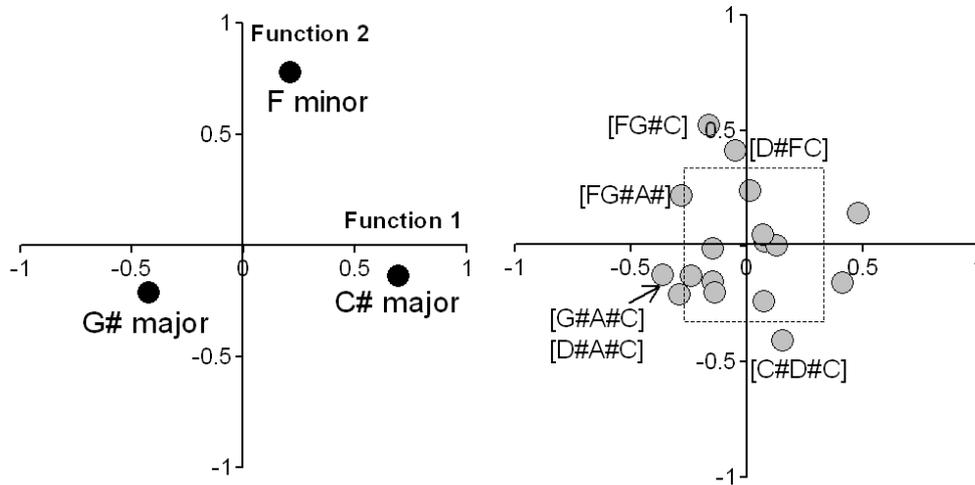


Figure 7. Experiment 2: Results of discriminant analyses for Stage 3

Black cycles indicate dependent variables (left), and gray cycles indicate independent variables (right). The dotted box indicates the boundary of $\pm .30$ of a structure coefficient.

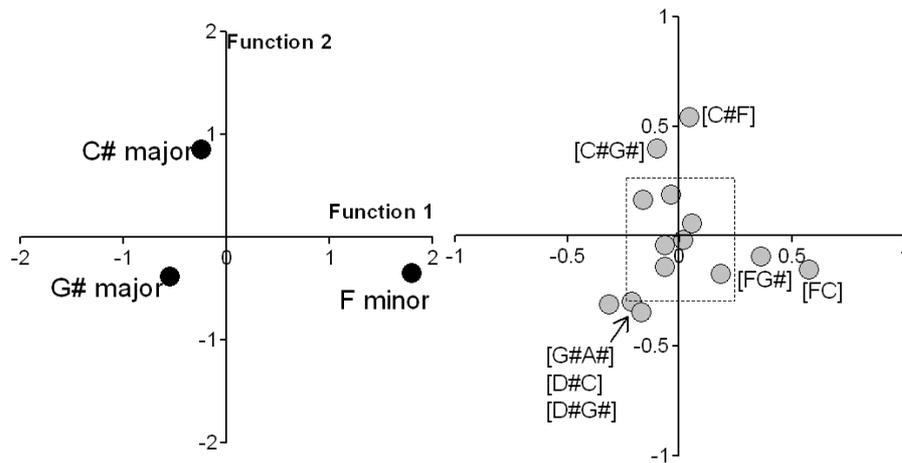


Figure 8. Experiment 2: Results of discriminant analyses for Stage 2

Black cycles indicate dependent variables (left), and gray cycles indicate independent variables (right). The dotted box indicates the boundary of $\pm .30$ of a structure coefficient.

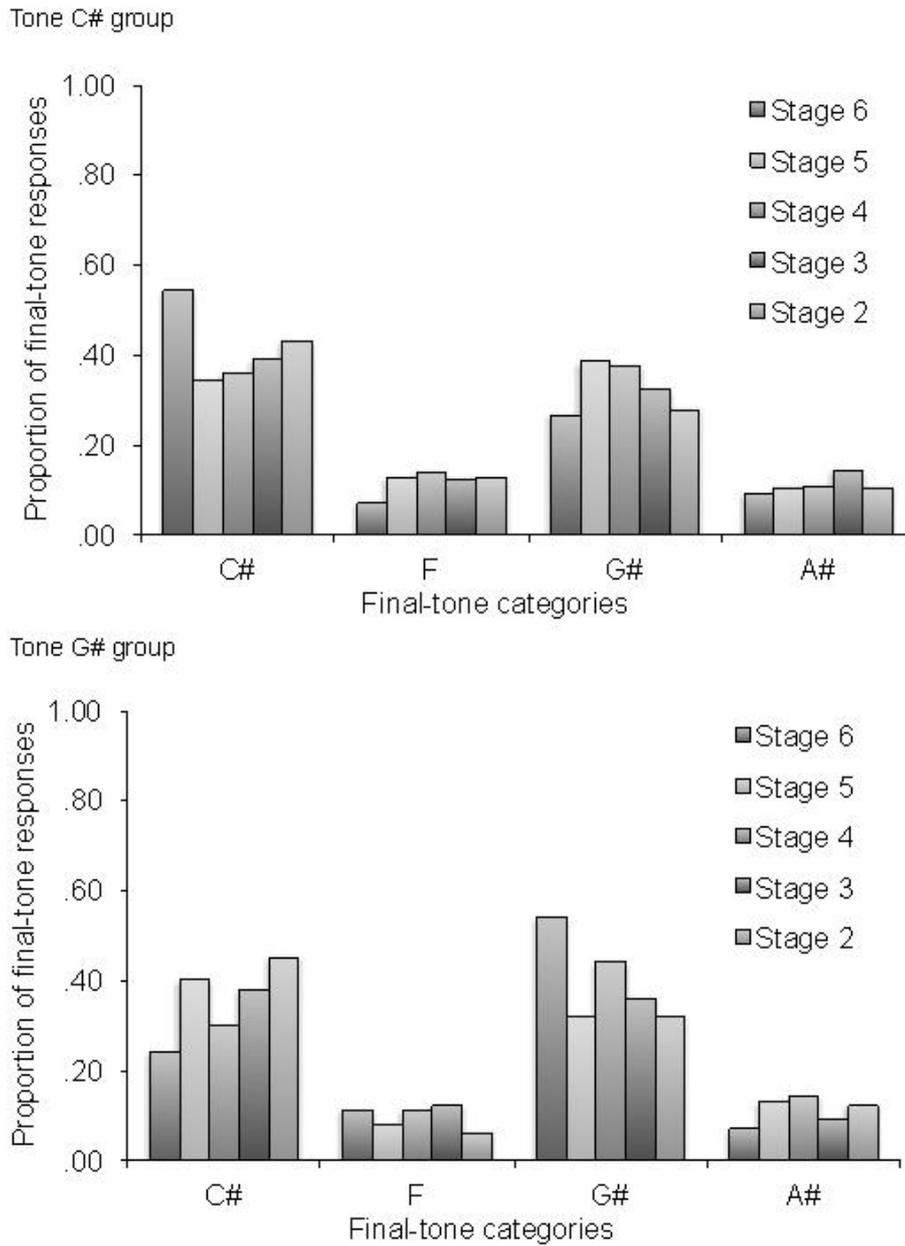


Figure 9. Distributions of final-tone responses for each of Tone C# and Tone G#

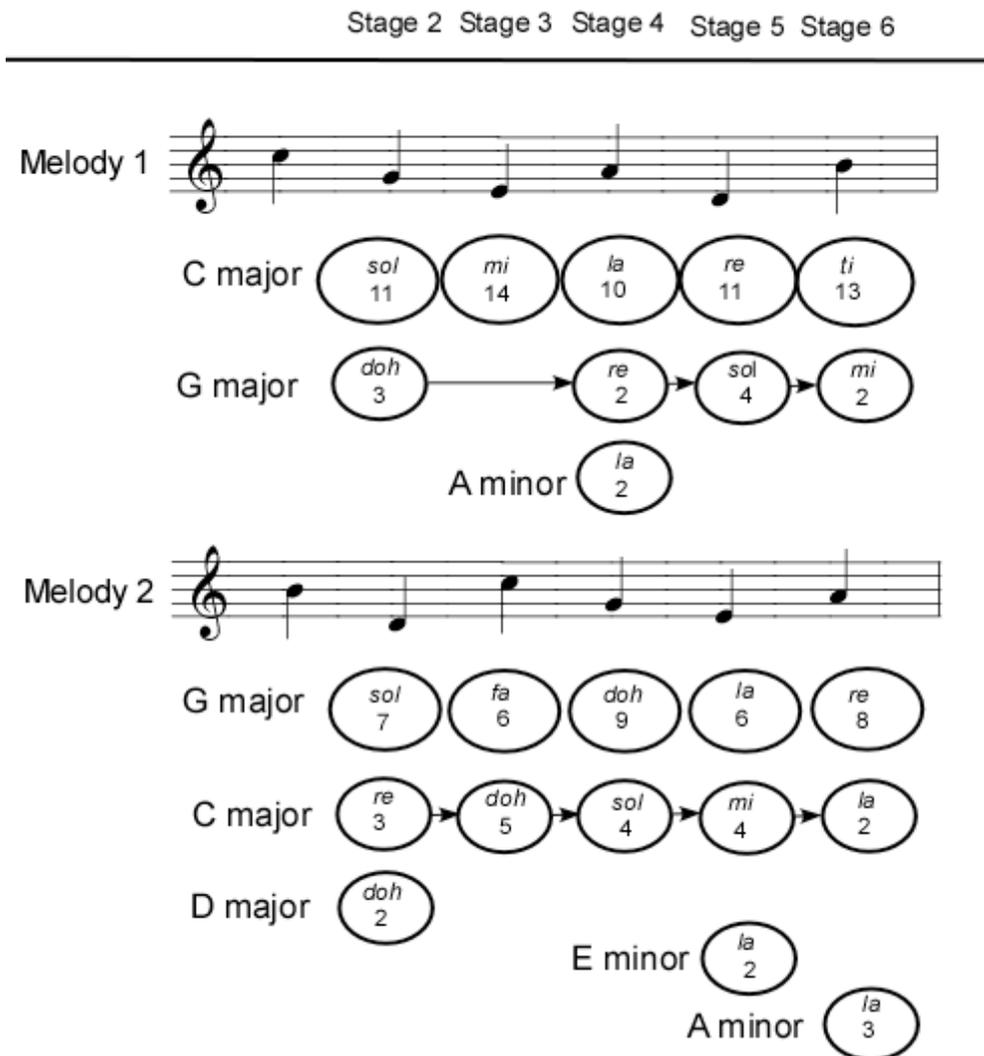


Figure 10. Participants' key responses for Melody 1 and Melody 2 (Experiment 1)

Ovals reflect key responses by two or more participants at each stage. Tone names in ovals are denoted in terms of 'movable doh'. The numbers in ovals show the proportions of participants who judged the key.