Pitch-distributional effects on the perception of tonality

Nicholas A. Smith
Mark A. Schmuckler

Division of Life Sciences, University of Toronto at Scarborough

Tonality is an important structural property of music, and has been described by music theorists (Meyer, 1956; Lerdahl, 1988) and psychologists (Krumhansl, 1990) as a hierarchical ordering of the pitches of the chromatic scale such that these notes are perceived in relation to one central and stable pitch, the tonic. This hierarchical structure is manifest in listeners’ perceptions of the stability of pitches in tonal contexts. For example, Krumhansl and Kessler (1982) had listeners rate how well each pitch of the chromatic scale fit into a variety of tonal contexts (e.g., scales, cadences) and found a consistent pattern of perceived stability; these ratings have come to be called a "standardized key profile." This profile, which appears in Figure 1 (with reference to a C major tonality) can be generalized to other keys by simply shifting the profile such that the tonic value is aligned with any of the twelve notes of the chromatic scale.
Given that the perceived stability of a pitch depends on the tonal context in which it occurs, it is important for musical processing that listeners apprehend the tonal structure of a musical passage. One approach to how listeners establish a sense of key (but see also Butler, 1989) stems from the recognition that, within a tonal context, those pitches that are perceived as most psychologically stable are also those pitches that are played most frequently, and for the greatest total durations; similarly, psychologically unstable pitches occur infrequently, and for the shortest durations. This observation has lead to models of tonality perception, such as the Krumhansl-Schmuckler key-finding algorithm (described in Krumhansl, 1990), that propose that listeners are sensitive to distributional information in music, and identify the key of a musical context based on the degree to which it matches with such acquired representations of tonal structures. In this vein, a number of studies (Coady, 1994; Laden, 1994; Oram & Cuddy, 1995) have demonstrated listeners’ sensitivity to distributional information.

The Processes of Tonality Perception

What psychological properties underlie listeners’ sensitivity to distributional information? One possibility is that sensitivity to this type of structure reflects two complementary processes: differentiation and organization. By differentiation we mean the distinguishing of pitches from one another in terms of some relevant dimension, such as their total duration or frequency of occurrence. In contrast, organization refers to a sensitivity to relations between differentiated pitches and the form in which the differentiated pitches are represented. In a series of experiments we examined these two processes by manipulating the distributional properties of random orderings of the chromatic scale.
Differentiation

The degree to which the pitches in the melodies were differentiated was manipulated by applying a power transformation to the standardized key profile values (Krumhansl & Kessler, 1982); the exponent for this power transform will be referred to as the *tonal magnitude* of the profile. This transformation has some interesting properties. First, changing the tonal magnitude has profound effects on the absolute differences between the values within the profiles. Thus, at a tonal magnitude of 0.5, the value for the 0th and 1st scale degrees are 2.52 and 1.49 (or .11 and .07 when expressed as proportions of the sum of all values), whereas at a tonal magnitude of 3.5 these same scale degrees have values of 645.22 and 16.56 (or .41 and .01) respectively. In contrast, varying the tonal magnitude of the profile only slightly influences the pattern of relative values in the profile (when expressed as standardized values). Thus, the correlation between the original and transformed profiles is quite high across a range of tonal magnitudes. Duration profiles for tonal magnitudes of 0.5 and 3.5 are shown in Figure 2.
Second, as the tonal magnitude approaches zero each value progresses toward the mean of all values, resulting in a profile that becomes flat, with the pitches less differentiated. In contrast, as the tonal magnitude increases the pitches become increasingly differentiated, although at very high tonal magnitudes the value for the tonic pitch so far exceeds the remaining pitches that the tonic dominates the profile, becoming the one and only pitch.

Organization

Organization was examined by either preserving or destroying the hierarchical structure present in the distributional information. In the hierarchical condition, the duration/frequency-of-occurrence values mirrored those of the standardized key profile, such that the longest (0th scale degree) and second longest (7th scale degree) pitches were a perfect fifth apart, and so forth. In contrast, in the nonhierarchical condition the duration/frequency-of-occurrence values were randomly assigned to pitches, thereby destroying the typical hierarchical relations between the pitches, while preserving the degree of differentiation among the pitches. If the perception of tonality reflects simple memory for longer or more frequently occurring pitches, then ratings of tonal stability should be similar in the hierarchical and nonhierarchical conditions. If, however, tonality perception requires hierarchical organization of distributional information, then ratings of the tonal stability of pitches will correspond to the duration/frequency of occurrence of pitches only in the hierarchical condition alone.

A third condition examined the possibility that listeners could extrapolate a hierarchical structure of tonality onto a set of pitches based solely on differentiating the tonic from all of the remaining pitches. In this binary condition the value for the tonic was identical to that of the corresponding value for the tonic in the hierarchical and nonhierarchical condition; in contrast, the remaining nontonic pitches all had the same value. If listeners can indeed extrapolate tonal structure by differentiating only the tonic, then ratings of tonal stability should demonstrate some (presumably hierarchical) differentiation among the nontonic pitch despite these pitches being undifferentiated in their durational properties. Sample hierarchical, nonhierarchical and binary duration profiles are shown in Figure 2.

Experiment 1

Experiment 1 examined whether listeners' perceptions of tonality were influenced by (a) the degree of differentiation (i.e., tonal magnitude) of the duration profiles, and (b) the presence or absence of hierarchical structure in these profiles.

Method

Participants.

Forty students at the University of Toronto at Scarborough, all meeting a 3-year minimum musical training requirement, participated in this experiment in exchange for payment or course credit. These listeners were assigned to one of two conditions: hierarchical and nonhierarchical. The mean years of musical training for listeners in each condition were 9.9 and 8.2 respectively; this difference in training was not statistically significant, \( t(19) = 1.38, p > .05 \).

Materials.

The stimuli consisted of a series of algorithmically composed melodies that were all 10 seconds in length and contained 24 notes, with each pitch of the chromatic scale occurring twice. The duration of each pitch was determined in the following way. For the hierarchical condition each value in the standard major key profile (Krumhansl & Kessler, 1982) was raised to one of ten tonal magnitude exponents (0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5). These transformed values were then expressed as a percentage of the sum for all 12 values, multiplied by 10000 (the duration of the melody in milliseconds) and divided by two (the number of times each pitch occurs in the melody). Melodies were then created by randomly ordering the 24 notes, with the onset of a note immediately following the offset of the previous note. For the nonhierarchical condition, a randomized version of the standardized key profile was created by randomly assigning duration values to the different pitches; all other aspects of stimulus generation were the same as in the hierarchical condition. Thus, hierarchical and nonhierarchical melodies contained the same number of long and short pitches, but differed in how these pitches were organized.
All melodies were played on a Yamaha DX7 synthesizer, set to an electric piano timbre. The synthesizer was connected to a Macintosh 8100 AV computer via a MIDI interface, and was controlled by a program written in the MAX programming language. Audio output from the synthesizer was fed into a Mackie 1202 mixer, and was presented to listeners through two Boss MA-12 micro monitors at a comfortable listening level.

Design and Procedure

The study employed the probe-tone method developed by Krumhansl & Shepard (1979). Each trial consisted of a presentation of a melody, followed by a 1 s silent interval, and a 2 s probe tone. The probe tone was one of the 12 pitches of the chromatic scale, and was played with the same timbre, pitch height (i.e., octave) and loudness as the melody. After each probe tone listeners rated on a 7-point scale the psychological stability of the probe tone, or how well they felt the probe tone fit into the context of the melody they heard.

Each block of the experiment contained 14 trials, with the same melody occurring on all trials within a given block; a different random order for the melody notes was employed for all listeners, however. The first two trials in the block were considered practice, and were intended to familiarize the listeners with the melody for that block. Trials 3 to 14 contained the 12 probe pitches presented in different random orders for all listeners. The entire experiment consisted of ten blocks of trials, corresponding to the ten tonal magnitudes tested. The order of the blocks of trials (i.e., different tonal magnitudes) was randomized across listeners. To avoid carry-over effects between blocks of trials, the melodies were randomly transposed to a different tonic in each block. The entire experimental session lasted approximately 45 minutes, after which each listener filled out a subject information form, and was debriefed.

Results and Discussion

Because the melodies were randomly transposed to different keys, probe-tone ratings reflect a goodness-of-fit measure relative to a given scale degree (e.g., 0th or 1st), rather than to an absolute pitch (e.g., C4 or C#4), for expository purposes, however, probe-tone ratings are presented here with reference to the key of C major. To facilitate comparison of the hierarchical and nonhierarchical conditions, the probe-tone ratings for the nonhierarchical condition were re-organized in terms of the scale degree of that pitch's duration. For example, the rating for the probe tone whose pitch had the longest duration in the melody was assigned to the pitch C. The rating for the probe tone whose pitch had the second longest duration was assigned to the pitch G, and so forth. Thus, the hierarchical organization of pitch durations destroyed by the nonhierarchical presentation has been reconstructed for purposes of data presentation and analysis.

A preliminary analysis of the data investigated the degree to which listeners' probe-tone ratings were intercorrelated at each level of tonal magnitude in the hierarchical and nonhierarchical conditions. Towards this end, a mean intersubject correlation for each listener (e.g., average correlation for listener 1 with listener 2, listener 3, and so on) was calculated, with these mean correlations analyzed in a two-way Analysis of Variance (ANOVA) with the within-subject factor of tonal magnitude (0.0, 0.5, ... 4.5) and the between-subject factor of condition (hierarchical versus nonhierarchical). Intersubject correlations varied as a function of both factors, with significant main effects for condition, $F_{(1,38)} = 70.00, \ p < .001$, and tonal magnitude, $F_{(9,342)} = 15.61, \ p < .001$, as well as a significant interaction effect between condition and tonal magnitude, $F_{(9,342)} = 10.19, \ p < .001$. Generally, the mean intercorrelation increased with increasing tonal magnitude, but only in the hierarchical condition. This finding suggests that in the hierarchical condition at high tonal magnitudes there is a common structure available to listeners that allows them to respond similarly.

The next step in the analysis involved examining whether listeners' ratings reflected the increased pitch differentiation produced by increasing tonal magnitudes, using a series of one-way ANOVAs, with the factor of pitch class. In the hierarchical condition, no significant differences were found in the ratings given to different probe-tones at tonal magnitudes 0.0 and 0.5 (both $p's > .05$). However, there were significant differences at tonal magnitudes of 1.0 and above, suggesting that listeners were differentiating between probe-tones at these tonal magnitudes. In the nonhierarchical condition, no significant differences were found in probe-tone ratings for the different pitches at any tonal magnitude.

Finally, the degree of organization of the differentiated pitches was examined by creating mean probe-tone ratings at each level of tonal magnitude, for hierarchical and nonhierarchical conditions separately, and correlating these probe-tone profiles with the standardized key profile of Krumhansl & Kessler (1982). Figure 3 shows the results of this analysis as a function of tonal magnitude and condition. In the hierarchical condition, correlations increased as a function of tonal magnitude, with high tonal magnitudes producing strong correlations. Thus, as tonal magnitude increased, probe-tone ratings more closely matched the standard description of tonal structure. In contrast, for the nonhierarchical condition there was no increase in correlations as a function of pitch differentiation.
Figure 3: Correlations between hierarchical and nonhierarchical probe-tone ratings and the standardized key profile, as a function of tonal magnitude.

In summary, Experiment 1 found that perceptions of tonality were affected by both the differentiation and organization of pitches. In terms of differentiation, the fact that a minimum level of tonal magnitude was required demonstrates that absolute, rather than relative, differences in pitch durations are important for perceiving tonality. In terms of organization, the finding that listeners only perceive tonality when the differentiated pitches are organized in a hierarchical fashion shows that differentiation alone is not sufficient, but that the relations between differentiated pitches are important.

Because listeners perceived tonality in melodies in which the pitches were differentiated by their total duration, but completely undifferentiated in terms of their frequency of occurrence (each pitch occurred twice in each melody), the present findings demonstrate that duration differences can be a crucial distributional property of music for tonality perception. It is an open question, however, whether or not frequency-of-occurrence information can also play as significant a role in perceiving tonality; this issue is addressed in Experiment 2.

Experiment 2
Experiment 1 demonstrated that hierarchical organization of pitches differentiated on the basis of note duration leads to a perception of tonality. Along with duration, however, frequency-of-occurrence information is typically differentiated hierarchically, such that duration and frequency of occurrence are naturally correlated. Interestingly, experiments that have teased apart the influences of duration and frequency of occurrence (Lantz & Cuddy, 1998) have observed that duration, but not frequency of occurrence, typically leads to a percept of tonality. A similar result is indirectly suggested by the findings of Experiment 1, in which perceptions of tonal stability were related to varying note duration of pitches, with frequency-of-occurrence information held constant. One way of interpreting this result is that undifferentiated frequency-of-occurrence information (which is the case when this variable is held constant, as in Experiment 1) does not counteract the impact of differentiated duration information. Unfortunately, such an interpretation says nothing about whether or not differentiation of frequency-of-occurrence information in and of itself can produce a percept of tonality. Investigating this question was the goal of Experiment 2.

Method

Participants.

Forty students at the University of Toronto at Scarborough participated in this experiment in exchange for course credit or cash. These listeners were assigned to either the hierarchical or non-hierarchical conditions. Listeners in the hierarchical and nonhierarchical conditions had an average 7.8 and 8.8 years of musical training respectively; this difference was not statistically significant, t(35) = 0.99, p > .05.

Materials.

Melodies were created that systematically varied the frequency of occurrence and total duration of their constituent pitches. All melodies were 10 s in length, and contained 100 notes. This study employed a more limited range of tonal magnitudes (0.5, 1.0, 1.5, 2.0 or 2.5), with the transformed values for each pitch expressed as a percentage of the sum of the transformed values, and with this percentage then determining the proportion of the 100 notes assigned to a given pitch. Thus, frequency-of-occurrence information was able to reflect variation in tonal magnitude.

The actual note duration for each pitch was calculated in one of two ways. In the uncontrolled duration condition all notes were sounded for 100 ms regardless of their pitch. As such, the total duration for each pitch increased with frequency of occurrence, providing a natural correlation between the two. In the controlled duration condition, the total duration for each pitch was set at 833 ms, with the duration of each individual occurrence of that pitch determined by dividing this total duration by its frequency of occurrence. Thus, the total duration of each pitch is held constant, thereby breaking the natural correlation between duration and frequency-of-occurrence information.

As in Experiment 1, melodies were either hierarchical or nonhierarchical, with the association between the frequency-of-occurrence values and pitch either preserving or destroying the hierarchical structure of tonality. In other respects the procedure was identical to that of Experiment 1.

Results and Discussion

Preliminary analyses examined the extent to which listeners' ratings were intercorrelated, using the same procedure as in Experiment 1. A three-way ANOVA on the average intersubject correlations was calculated, with the within-subject factors of tonal magnitude (0.5, 1.0, 1.5, 2.0, 2.5) and duration type (controlled versus uncontrolled), and the between-subject factor of condition (hierarchical versus nonhierarchical). This analysis revealed a main effect of condition, F(1,38) = 6.16, p < .05, with listeners in the hierarchical condition showing significantly more agreement than those in the nonhierarchical condition. There were also main effects of duration type, F(1,38) = 35.96, p < .001, with listeners' ratings showing more agreement in the controlled duration than in the uncontrolled duration condition, and tonal magnitude, F(4,152) = 8.08, p < .001, with more intersubject agreement in higher tonal magnitude conditions. All 3 two-way interactions were significant: condition by duration type, F(1,38) = 0.44, p > .05, condition by tonal magnitude, F(4,152) = 5.05, p < .001, and tonal magnitude by duration type, F(4, 152) = 4.47, p < .01. However, the three-way interaction effect of tonal magnitude, duration type and condition was not significant, F(4, 152) = 1.96, p > .05. These results suggest that a commonly perceptible structure is more readily available to listeners in the hierarchical condition, at higher tonal magnitudes, and when duration is uncontrolled.

Pitch differentiation was again examined in a series of one-way ANOVAs, with the factor of pitch class. These analyses found significant differences in probe-tone ratings for the uncontrolled hierarchical and nonhierarchical conditions, suggesting that differentiation was a function of pitch duration and not frequency of occurrence. In addition, the strength of this effect increased with increasing tonal magnitude, suggesting that listeners were differentiating pitches on the basis of their absolute, rather than relative, duration differences.
Finally, the degree of organization was examined by correlating mean probe-tone profiles for the various conditions with the standardized key profile of Krumhansl and Kessler (1982); the results of these correlations are shown in Figure 4. For all conditions the strength of the correlation increased as a function of tonal magnitude. However, the correlation only reached statistical significance in the hierarchical uncontrolled duration condition, suggesting that the perception of tonal structure is especially dependent on hierarchically organized absolute differences in duration.

**Figure 4**: Correlations between hierarchical and nonhierarchical probe-tone ratings and the standardized key profile, as a function of tonal magnitude and duration type.

In summary, Experiment 2 confirmed the finding that listeners' perceptions of tonality were affected by both the degree to which pitches were differentiated and the way in which the differentiated pitches were organized. Adding to these findings is the result that perceptions of tonal stability were primarily dependent upon increasing duration information, and were not driven by frequency of occurrence per se. When duration information was held constant (i.e., the controlled duration condition), listeners showed some sensitivity to pitch differentiation through the general increase in correlation with the standardized key profile, although not enough sensitivity to recover the typical sense of tonal organization.

**Experiment 3**

Although the previous studies suggest that it is the hierarchical organization of the complete set of differentiated pitches that leads to a percept of tonality, it might be that less severe forms of hierarchical structure can similarly drive tonal perception. Is it, for example, necessary that there be multiple levels of differentiation and organization among pitches, or can the percept of tonality be instantiated with a minimum amount of such organization? This question can be examined by creating a new type of melody whose duration profile has only two possible levels of differentiation. In these binary profiles the duration value for the tonic is preserved, with the durations of the remaining pitches set to a different level; thus, the only differentiation is between tonic and nontonic pitches. With these binary profiles, the tonal magnitude manipulation thus varies the ratio between...
the duration of the tonic and the remaining pitches.

Method

Participants.

Twenty students at the University of Toronto participated in this experiment in exchange for payment or course credit. They all met the 3-year minimum musical training prerequisite, having an average 8.15 years of training.

Materials.

The melodies were identical to the hierarchical melodies of Experiments 1 and 2, except that the duration values for the 11 nontonic pitches were set to the same value. Tonal magnitude values varied as in Experiment 1, from 0.0 to 4.5, in 0.5 increments.

Procedure.

The procedure was identical to that of Experiments 1 and 2.

Results and Discussion

A preliminary analysis examined the degree of intersubject correlation in the probe-tone ratings using a one-way ANOVA with the factor of tonal magnitude. Overall, there was a general increase in the average intercorrelation (from -.02 to .12), indicating increased levels of agreement between subjects as a function of tonal magnitude; these differences were significant $F(9, 171) = 7.51, p < .001$.

The analysis of differentiation in listeners’ probe-tone ratings, at each level of tonal magnitude, revealed significant differences between ratings of different pitch classes at tonal magnitudes of 2.5 and higher. To determine whether these effects were due to the differentiation of the tonic from nontonic pitches, or from differentiation among the nontonic pitches themselves, a second set of ANOVAs was performed on the nontonic pitches alone. Although significant differences were found at tonal magnitudes of 2.5 and 3.0, there were no differences in probe-tone ratings at higher tonal magnitudes.

Finally, mean probe-tone profiles at each level of tonal magnitude were correlated with the standardized tonal hierarchy ratings of Krumhansl and Kessler (1982). As shown in Figure 5, a general increase in the strength of the correlation was found as a function of increasing tonal magnitude, reaching statistically significant levels at tonal magnitudes of 4.0 and 4.5. To examine the possibility that these correlations were exaggerated by a high leverage tonic value, the correlations were recalculated for nontonic pitches only. Dropping the tonic from the analysis eliminated the trend, lowering all correlation coefficients to nonsignificant levels.

In summary, Experiment 3 demonstrated that the differentiation of the tonic from nontonic pitches alone was not sufficient to evoke perceptions of tonality, and that further differentiation among the nontonic pitches, as occurred in the hierarchical condition of Experiment 1, is necessary. One open issue with these findings involves the degree of differentiation, or the number of hierarchical levels, necessary to invoke a percept of tonality. Although this question has been at least indirectly considered in some previous work (e.g., Takeuchi, 1994; Temperley, 1999), a thorough investigation of this issue has yet to be undertaken. Regardless, the suggestion here is that, at the very least, multiple hierarchic levels are required for adequate tonal perception. Interestingly, such findings have implications for models of musical key finding (e.g., the Krumhansl-Schmuckler key-finding algorithm, in Krumhansl, 1990), in which tonality is assessed via a pattern-matching process with the standardized key profile values.

General Discussion

These experiments illustrate the importance of distributional information as an indicator of tonality in music, and that listeners’ sensitivity to this information reflects both the differentiation and organization of pitches. The consistent effect of tonal magnitude on perceptions of tonality suggests that listeners differentiate pitches on the basis of absolute, rather than relative, differences in the total duration of pitches. Comparisons of hierarchical, nonhierarchical and binary melodies suggests that hierarchical organization of duration information is similarly crucial to the perception of tonal structure. Furthermore, independent manipulation of the total duration and frequency of occurrence of pitches lends further support to the claim that duration plays the primary role in pitch differentiation.
A further finding in these experiments is that random orderings of notes can invoke perceptions of tonality in listeners. This result contrasts, for example, with the work of West and Fryer (1990), who failed to find tonality percepts when they asked listeners to provide tonality ratings after presenting random orderings of a set of pitches that uniquely specifies a given key. One likely reason for this difference in findings is that West and Fryer failed to provide any differentiation among the pitches in their melodies, employing essentially a single level hierarchical organization (presence versus absence of various pitches). Given the importance of both pitch differentiation and multiple hierarchical levels demonstrated by the current experiments, the failure of West and Fryer's stimuli to produce tonal percepts is understandable.

The present experiments also somewhat undermine the importance of rare intervals in the perception of tonality (Butler, 1989). Given that these melodies were chromatic, the rare interval information presumably critical for perceiving tonality was rendered highly ambiguous; nevertheless, listeners were able to perceive the tonal structure inherent in these stimuli. Thus, theories of tonality perception that argue that listeners identify the tonality of music on the basis of rare intervals would have a difficult time accounting for the current findings.

The finding that the tonal stability of a pitch is related to its proportion of the total duration may reflect a very general principal of auditory pattern processing, and fits well with the proportion-of-the-total-duration rule (Kidd & Watson, 1992) in which selective attention to an element in an auditory pattern is a function of the proportion of the total duration of the pattern for which that element is present. This finding, among others, has been incorporated into Lufti's (1993) component-relative-entropy (CoRE) model of auditory pattern analysis that proposes that the discriminability of an element in a pattern is a function of the degree to which its duration (or other dimension) contributes to the variance in durations among all elements. Increasing the tonal magnitude of a duration profile has precisely the effect of increasing the tonic duration's proportion of the total duration, as well as its contribution to the variance of the duration of all pitches.

References


