‘Tonal’ vs ‘atonal’: Perception and tonal hierarchies

Robertas Budrys
Lithuanian Academy of Music and Theatre, Vilnius, Lithuania
budrys at super.lt

Rytis Ambrazevičius
Faculty of the Humanities, Kaunas University of Technology, Kaunas, Lithuania
Lithuanian Academy of Music and Theatre, Vilnius, Lithuania
rytisam at delfi.lt - http://www.rytisambrazevicius.hmf.ktu.lt

Proceedings of the fourth Conference on Interdisciplinary Musicology (CIM08)

Background in music theory. Many factors have formed quite a few definitions of tonality (Hyer, 2001). For instance, should the term ‘tonal’ be applied only to the Western music of common practice period, or rather it is applicable to other musics as well? Most theoreticians define tonality, in general, as systematic organization of pitches where at least one pitch class is emphasized. Notion of atonality, obviously, covers all music that is not tonal (Lansky, Perle, and Headlam, 2001). In a frame of 12ET, twelve-tone serial music with pitch classes of equal significance should be the most opposite to tonal music (Schoenberg, 1975). Provided this discourse is satisfactory, two questions still remain open: first, to what degree the types of tonality can be differentiated (as in Kholopov, 2003), second, to what degree the theoretical discourse on tonality/atonality reflects in cognition?

Background in music cognition. Well-known Krumhansl’s theory of tonal hierarchies (1990) describes perception of tonality as operating tonal profiles formed by exposure of a listener to a certain soundscape and revealed by probe-tone technique. Whereas this model is quite clear in the case of tonal music, it raises some questions in the case of atonality. Krumhansl proposed that listeners, when perceiving tonal structures in atonal music, apply two distinct strategies depending on their musical training: ‘normal’ or ‘reversed’. However, she only employed the tone series used by Shoenberg and Webern. Lerdahl (1989) proposed that the contextual salience cues play the most important role in the perception of atonal music, i.e., the listening rules differ from those in tonal music as atonal ‘pitch space is flat’. Gregory (1997) tried to find whether the tonal hierarchies perceived in two pieces by Shoenberg and Webern ‘were correlated to the frequency of occurrence and total sounding duration’.

Aims. We aim to reveal the regularities of estimation of musical excerpts in terms of perceived tonality/atonality and to collate these results with the perceived tonal hierarchies according probe-tone tests.

Main contribution. In Experiment 1, 90 subjects with mixed musical experience (mostly music students) listened to 50 recorded excerpts of real music performances; the durations of the excerpts varied from 3 to 14 s. The subjects were asked to rate perceived tonality/atonality of the excerpts in five-point scale. The tasks were preceded by normalization tests. The perception of the tonality/atonality continuum showed features characteristic for categorization: two categories ('tonal' and 'atonal') were perceived. The categorization was more distinct in the case of the longer excerpts. The ratings of the longer excerpts also correlated better with the general evaluations based on music theory, but they are hardly consistent with the typologies of tonality.

In Experiment 2, 8 chord progressions (4 in major and 4 in minor keys) and 10 excerpts of the real music performances (selected from Experiment 1) were presented to 30–45 subjects. The subjects were asked to rate how well the tones fitted to the progression or excerpt in the probe tone task (Krumhansl, 1990). The main difference from Krumhansl’s experiments was the broad spectrum of real music performances (the composers ranged from Mozart to Webern). When listening to the music excerpts evaluated in Experiment 1 as quite tonal, the listeners perceived quite distinct tonal hierarchies. The evaluations of the music excerpts perceived as atonal were accidental, i.e., the correlation of the individual evaluations was faint. In this case, the tonal hierarchy profiles were flat and high, in comparison with the case of tonal music. The correlation between the perceived tonality/atonality (Experiment 1) and the means of tonal hierarchy profiles (Experiment 2) was significant (91 %).

In summary, perception of tonality/atonality is categorical, yet hardly reflects the typological peculiarities. The parameters of tonal hierarchies according probe-tone tests correlate with perceived tonality/atonality. The results are obtained employing considerable number of pieces and listeners.

Implications. A listener, unlike theoretician or composer, often applies relatively simple strategies in the design of pitch structures or/and strategies not consistent with the theoretical or compositional models. Therefore music theory and composition courses should take into account phenomena of tonality/atonality perception. 'The best music arises from an alliance of a compositional grammar with the listening grammar’ (Lerdahl, 1992).

The present study could be further extended to model cognitive phenomena of tonality/atonality.
A comprehensive account of the phenomena of tonality/atonality requires expertise in both the humanities (music history, theory, analysis) and the sciences (music psychology, statistics).

Often, when discussing the phenomena of tonality and atonality, various theoretical intermediate types are differentiated (Kholopov, 2003; etc.). The present study deals rather with an ‘immediate’ perception of tonality/atonality (presumably not including any theoretical speculations). Two experiments were conducted to study the perception of tonality/atonality and the tonal hierarchies. Then the results of the experiments were collated.

The strategy of the study was developed and the results were discussed by both Robertas Budrys and Rytis Ambrazevičius. The experiments were prepared and done, and the succeeding calculations were made by Robertas Budrys.

**Experiment 1**

**Preparation**

**Hypothesis.** The tonality / atonality (T/A) of real musical pieces as particular quality of music should be perceived similarly by most subjects featuring similar musical backgrounds and living in certain soundscape. The experiment was designed to verify this statement and to reveal the principles of T/A perception.

**Samples.** 50 pieces of diverse tonal structure were selected; from Mozart to Webern. Mostly works of the first half of the 20th century were considered. Three excerpts with length of 3, 7, and 14 s were taken from the same place of each piece. The excerpts had to conform to the following requirements: twelve-tone equal temperament, similar tempo, polyphony, instrumental timbre, and homogeneity of the musical material.

All the excerpts were taken from audio recordings of good quality. The excerpts of the same length were combined into three sets (each of 50 excerpts). Also they were faded in and out (fade-ins and fade-outs of 0.5, 1 and 2 s, respectively) to reduce the effect of the first and the last sounds. The same random order was applied for the excerpts in all three sets, but with avoidance of the neighboring excerpts with possibly similar T/A. The excerpts were separated by breaks of 10 s to avoid the possible influences of the previous excerpts.

**Normalization samples.** Two musical samples were prepared to normalize the evaluations of T/A, each of them representing an extreme tonality and an extreme atonality. The first one was the nursery song *Twinkle, Twinkle, Little Star* (Fig. 1), and the second one was the two-part twelve-tone composition made of four rows that were randomly generated from notes of equal duration (Fig. 2). Both samples were 14 s in length, with fades in and out. They were generated applying software *Sibelius*, and piano timbre was assigned.

![Figure 1. Normalization sample 1: Twinkle, Twinkle, Little Star (fragment).](image1)

![Figure 2. Normalization sample 2: Two-part twelve-tone composition (fragment).](image2)

**Participants.** The procedure was applied to three groups of subjects. All the subjects were musicians: students and teachers of performance and musicology, and teachers of children music school. Group 1 included 31 subjects (it was presented with 3 s excerpts), Group 2 included also 31 subjects (7 s excerpts), and Group 3 included 32 subjects (14 s excerpts). The total number of the filled questionnaires was 94, four of them were rejected (the subjects misunderstood the task).

**Procedure**

In the beginning, the subjects heard the two normalization samples to define the concepts of tonality and atonality. Then they heard 50
excerpts of equal lengths. The subjects had to rate their sense of T/A of each excerpt in 5 point scale, 1 for extreme tonality, and 5 for extreme atonality. The subjects were asked to ignore their theoretical knowledge that could be called up by identification of the style, the composer, or even the particular piece, and to rate only the spontaneous sense of T/A.

Results
The results of each group were processed separately. T/A rating averages and standard deviations were calculated for each excerpt. Fig. 3 shows ratings of Group 2 (7 s excerpts). The tendency can be observed that the standard deviation of rating is largest in the middle of the rating scale (around 3 points), while it decreases towards the point extremes (1 and 5 points).

The results of Groups 1 and 3 are similar to those of Group 2. The pair correlations between the groups are 96-97% (for the rating means) and range from 77% to 86% (for standard deviations of the ratings). The standard deviations tend to decrease with the increasing duration of the excerpts: the mean standard deviation (averaged across all pieces and all participants in the certain group) equals 0.84 for Group 1 (3 s excerpts), 0.82 for Group 2 (7 s), and 0.77 for Group 3 (14 s). Also it seems that the longer the excerpt, the better the T/A ratings correspond to the evaluations of the traditional music theory.

Experiment 2
Preparation
Hypothesis. According to the Krumhansl theory of tonal hierarchies (TH), music listeners perceive pitches through certain cognitive system (Krumhansl, 1990, 9). One can assume that the stable THs are created by tonal music and, vice versa, the unstable THs are created or are not created at all by atonal music. The experiment based on Krumhansl’s probe tone method (Krumhansl, 1990, 21, 25–27) was designed to verify this statement.

Samples: musical pieces. 10 pieces of diverse tonal structures were selected from the 50 pieces used in Experiment 1. Single excerpt was taken from each piece, from the same place of the piece as in Experiment 1. The excerpts varied in length (but they were no longer than 14 s) and presented complete musical patterns which had no obvious transitions and modulations. The selection and processing of the excerpts was based on the methodology and conclusions of Experiment 1. The different excerpts were combined into three sets (of 3, 3, and 4 excerpts). The main difference from Krumhansl’s experiments was the broad spectrum of real music performances.

Samples: chord progressions. In addition, four chord progressions were created: IV–V–I, ii₆–V–I, vi–V–I, and I₆–IV–I₆–V–I. By change of tonal centers and modes (natural major and harmonic minor), 24 variations of these progressions were made and combined into six sets (4 different progressions; each in different key; two of major and two of minor modes). All the progressions were generated applying software Sibelius, and piano timbre was assigned.

Participants. The procedure was applied to 6 groups of subjects; 15 subjects in each group. All the subjects were musicians: students and teachers of performance, musicology, composition and music technologies. The experiment was designed so that each chord progression was heard by 15 subjects, and each excerpt was heard by 30 subjects. The total number of the filled questionnaires was 90.
Procedure

The subjects heard four chord progressions and three or four excerpts of pieces, all prepared according to the probe tone technique: the context was repeated 12 times and followed by a different probe tone (one of 12 pitch classes) each time. The probe tones were made from Shepard tones (length 0.8 s). The subjects had to rate how well the probe tone fitted to the context in a musical sense by giving points from 1 to 7 (1 for ‘no fit’ and 7 for ‘excellent fit’). The subjects were asked to ignore their theoretical knowledge, not to attempt to remember did the probe tone correspond to any sound of the context just heard, and to use the rating scale as wide and gradual as possible. The procedure was applied to six groups of subjects. The different sets of chord progressions and excerpts were presented to each group.

Results

The TH profiles were calculated for each chord progression and excerpt. The profiles of progressions were transposed to C major and C minor keys and the results of the same progressions were merged (8 profiles were obtained; four of major progressions and four of minor progressions), because the same progressions in the different keys were rated quite similarly (the correlation between them was found to be from 83 to 96%).

The generalized major/minor profiles were formed by merging all THs of major/minor progressions into two THs. As side result, the comparison of the major and minor profiles with the profiles obtained by Krumhansl (1990, 30) is presented (Figs. 4 and 5). The results of our experiment correspond to Krumhansl’s results quite well: the correlation is 98% and 96% for major and minor profiles, respectively. Some differences, however, result (seemingly, first of all) from the larger number of the participants in our case and, possibly, from the nuances of the different perception. Also some influence of the musical material used can be suspected: we applied simple piano chords instead of Shepard chords in the Krumhansl’s experiments.

The TH profiles of the excerpts rated as quite tonal in Experiment 1 are fairly distinctive, the pitches are clearly differentiated, and the pitch structure of the score is identifiable. The TH profiles of the atonal excerpts are indistinctive, most of the pitches are rated similarly and as quite stable (they get high points, on the average), and the pitch structure of the score is unidentifiable (Fig. 6).

Figure 4. Comparison of our and Krumhansl’s (1990) TH profiles for major progressions.

Figure 5. Comparison of our and Krumhansl’s (1990) TH profiles for minor progressions.

Figure 6. TH profiles for three pieces (see Fig. 3). The pitches are transposed so that C gets maximum rating.
Two basic statistical values were calculated for each TH of the chord progressions and excerpts: the mean and the standard deviation. TH mean shows how stable all the pitches of the musical material are perceived, on the average. TH standard deviation shows how distinctive the TH profile of the musical material is. The correlation matrix was designed for each chord progression and excerpt to study the similarity of individual ratings (individually perceived THs). TH means and standard deviations of the samples show that the listeners perceive the pitches of tonal music (according to the results of Experiment 1) as quite unstable (on the average) and differentiated, and, vice versa, the pitches of atonal music are perceived as quite stable and get similar points. The correlation of the individual ratings of the tonal excerpts and chord progressions ranges from 32 to 55%, whereas the correlation is only from 1 to 4% in the case of extremely atonal excerpts. Thus the less tonal (the more atonal) the musical material, the more accidental the individual ratings.

Discussion

As the results of Experiment 1 show, the T/A perception is of categorical nature. When evaluating T/A of music piece (its excerpt) the listeners are able to separate two categories, i.e., the tonal and atonal poles. They give relatively similar points to the extremely tonal pieces as well as to the extremely atonal pieces. The border between these categories is marked by the relative uncertainty of the estimations. It should be stressed, however, that the differentiation between the categories is not as distinct as in the case of pitch classes, for instance. Anyway, only the two categories are found; there is impossible to detect any intermediate categories in the continuum from ‘tonal’ to ‘atonal’.

The comparison of the results obtained with different durations of the excerpts leads to the conclusion that the tendencies in T/A perception are influenced by the durations quite faintly. Nevertheless, some traits of the influence can be observed. The longer excerpts can be easier classified as tonal or atonal. They induce clearer and more defined response.

The results of Experiment 2 show that the TH principles and study methods suit tonal material better. When listening to atonal music the respondents do not perceive clear and unambiguous THs. Two standard modes of perception (TH-type and no TH-type) explain why we tend to separate two categories, i.e., ‘tonal’ and ‘atonal’. Nevertheless, not distinct THs of atonal pieces show the general tendency: all scale pitches are perceived as relatively stable; to be precise, they are perceived as more stable (on the average) than the pitches of tonal pieces. Thus means of TH profiles can be used for the evaluation of T/A. Indeed, the T/A rating means (across the respondents) and the means of TH profiles (across the respondents and the scale pitches) show strong correlation (Fig. 7) of 91%. Moreover, the T/A rating means also correlate negatively with the standard deviations in TH profiles (among the different scale pitches) quite well; -87% is found.

In summary, the phenomenon of the perception of tonal hierarchy determines directly the perceptual quality of ‘tonal/atonal’.

Figure 7. Dependence of T/A ratings on TH ratings averaged across the scale pitches. 10 pieces selected for Experiment 2 are considered.

Acknowledgments. We wish to thank all the participants of the experiments.

References

Gregory, N. (1997). The perception of tone hierarchies in atonal and 12-note music; two


\[\text{The first three progressions were used in the Krumhansl’s experiments (1990, 25).}\]