

On Cognition of Musical Grouping: Relationship Between the Listeners' Schema Type and Their Musical Preference

Mitsuyo Hashida^{1,2}, Kenzi Noike², Noriko Nagata³, and Haruhiro Katayose^{2,3}

¹ Graduate School of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama, Wakayama 640-8510, Japan
<http://www.sys.wakayama-u.ac.jp/>

² PRESTO/Japan Science and Technology Agency
<http://www.kyotyo.jst.go.jp/index.html>

³ Department of Science and Engineering, Kwansai Gakuin University, 2-1 Gakuen, Sanda, Hyogo, 669-1337, Japan
{hashida, noike, nagata, katayose}@ksc.kwansei.ac.jp
<http://www.m-use.net/research/listening/>

Abstract. We assume that there are various musical groupings of perceptions according to the degree of schemata and there are two dominant music grouping schemata; (a) **accent-oriented grouping schema** and (b) **phrasing schema** (musical expression referred to as the Rainbow type). In order to verify these hypotheses, we investigated how listeners' groupings change when the inner voice of Beethoven's Piano Sonata "Pathétique" was replaced with chords. We eventually succeeded in identifying three listening groups; those who have a strong (a) schema (type **A**), those whose (a) is prior to (b) (type **AF**), and those whose (b) is prior to (a) while paying attention to their inner voice (type **FA**). We verified that type **A** listeners prefer Rap music, Rock music, listening in a lively place, listening to party music, and listening to lyrics, while type **FA** listeners prefer Bach, Chopin, and listening alone and quietly.

1 Introduction

How do human beings listen to the music? How do human beings understand musical structure and perceive musical groups? These questions are of the greatest concern for music composers and performers. At the same time, they are essential questions for people studying what makes humans feel entertained.

The perception of music grouping is said to be a function of schemata. In perception, there are many organizing principles called gestalt laws, e.g., the laws of proximity and similarity [1]. A gestalt law says that we are innately driven to experience things in as good a gestalt as possible. Musical performances are thus understood in these terms as an art in which gestalt factors are elaborated ingeniously. However, it is not clear how these laws are unified to achieve music groupings. We assume that the wide variety of personal music groupings results from a difference in schemata strengths formulated in accordance with the listener's experience.

According to gestalt laws, the prominent note is perceived as the starting note of the music group, which consists of the note and the following contiguous notes. On the other hand, in the typical expression of phrasing, the most prominent note appears in the middle of the phrase. Without a doubt, both schemata must be valid in themselves, yet they are contradictory schemata for music grouping.

It is said that the primary goal of conductors and performers is to clarify the structure of the music to be played and to give expression to the music as her/his intention is understood by the audience [1]. This may lead to the view that there has to be a correct grouping, i.e., a correct way of listening to the music. In the research field of music psychology, there have been a few reports that focus on the variety of music group perceptions. However, in our analyses of a performance rendering contest [2] and a preliminary experiment on music listening [4, 5], we confirmed that even listeners who have a lot of musical experience do not have unique music groupings.

The paper addresses two points: 1) the analysis of grouping by subjects, and 2) the examination of background factors affecting grouping type. In Section 2, we explain typical performance expressions of music groups that are related with two grouping schemata, and explain the keywords and concepts of the experiments described in Section 3. Section 3 describes the experimental procedure to classify the grouping types of listeners. In Section 4, we discuss the musical preference background of each grouping type and summarize the subjects to be dealt with in the future.

2 Group Recognition of Music

A person perceives groups by listening to a stream of sound [6~12].

Music grouping is a fundamental process that listeners need to understand and enjoy music. In this section, we explain two typical schemata for music grouping. Next, we describe the preliminary investigation regarding group recognition and describe the sort of “attention” employed in the design of the experiments.

2.1 Schemata for Music Grouping

There are many perceptual organizing principles called gestalt laws; e.g., the law of proximity, law of similarity, and law of continuity. These principles are summarized in terms of music as follows:

Score level: grouping based on relative pitch-interval, direction of contour, and/or combination of these principles.

Signal level: grouping by IOI (inter onset interval), OOI (offset onset interval) and/or intensity level of adjacent notes.

These grouping principles are easy to understand, and most cognitive music theories adopt them in order to formulate grouping rules (e.g. GTTM [6]). However, the methods of how to set quantitative parameters for each rule and of conflict resolution between rules have not been formulated yet. The lack thereof has become one of the key issues of musical information science [13, 14]. Below, we introduce two principal schemata of music grouping.



Fig. 1. Perceptive Grouping Based on Gestalt

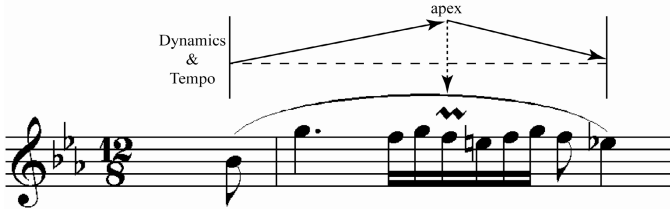


Fig. 2. Phrasing Expression

Accent-oriented grouping. Proximity is regarded as the most fundamental evidence for grouping music, and most grouping theories refer to it. In particular, Mozart's piano sonata K.331 (Fig. 1) is often quoted in explaining proximity. The explanation is as follows; the perceived groups are indicated by placing upper brackets if a small rest is inserted after every bar, and are indicated by placing lower brackets if a small rest is inserted just after quarter notes of the 1st and the 2nd bar.

Takeuchi measured performances of typical editions of the piece (Henle edition (upper) and Peters' edition (lower)) and showed that the group starting notes are played louder in both editions [15].

The principle of proximity and Takeuchi's suggestion that louder notes tend to be starting notes of groups (or performers give accent to the group starting notes) may be best regarded as independent concepts. However, as for proximity, it is rational to suppose that the greater the distance between adjacent events is, the more each event will be perceived as the prominent "accented" event. Thereby, combining the principle of proximity and Takeuchi's suggestion, we assumed a grouping schema by which a louder or leaped note after a long IOI or rest is perceived as a starting note of a group composed of successive notes. We refer to this schema as the **AOG schema** (accent-oriented grouping schema).

Phrasing. Another important clue for music grouping is the expression of phrases, referred to as phrasing. Expressing changes to tempi and dynamics by drawing an arc of a rainbow is the most widely known of the techniques to express phrasing (see Fig. 2) [2, 9, 16]. Although this phrasing technique may be a way to express a group in a range of a physical breath, it is also natural to think that this familiar expression might be used to formulate a principle schema for grouping music. Gestalt psychology informs us that phrasing can be regarded as an expression using the principle of proximity. In this paper, we refer to this grouping schema based on musical expression as the **phrasing schema**.

2.2 Preliminary Investigation

Here we introduce some points related to music grouping that we discovered during our preliminary investigation [5].

Variety of music group perceptions. We investigated how much performers' well-elaborated intention about music grouping is conveyed to listeners by conducting an experiment with 101 subjects comprising experienced listeners (more than 8 years experience in playing a musical instrument) and inexperienced listeners. Experienced listeners were able to understand intention better than inexperienced listeners ($p < 0.05$). On the other hand, the experiment also showed the number of the listeners of the experienced group who perceived fully the same grouping that the player intended was only half. The other perceived groupings could be classified into several types.

Influence of performance parameter. In order to elaborate on Takeuchi's finding in [15], we implemented a kind of morphing system, by which we can quantitatively investigate the influence of the operation to make a note prominent by giving more intensity and the operation of giving a rest to a music grouping. We carried out experiments (Fig. 1) on two subjects who had more than 20 years of musical experience (one has a masters degree in music education and the other, a masters degree in composition). The experimental results showed that the subjects' parameter distributions (intensity and rest) to judge the upper from the lower group (in Fig. 1) were significantly different.

Change of grouping by attention control. We tried to determine if the music grouping changes temporarily, when the listeners are given another attention focus than what they would usually focus on. Using the musical example described in section 3, we explained the phrasing of the accompaniment to listeners who were considered to pay attention to the melody. One third of listeners reported that their grouping unexpectedly changed from melody to accompaniment.

3 Listening Experiment

Our preliminary experiment suggests that music grouping is not always unique and that listeners can be classified according to a number of music groupings. We also verified that a listener's musical experience has some correlation with his/her music grouping and that attention control may influence a listener's grouping temporarily and selectively. Thus, the experiments were on classifying subjects according to grouping characteristics. We focused on the predominance of the **AOG schema** and **phrasing schema**. We have to take account of the fact that it is not easy for every subject to describe her/his own perception. Sometimes a perception may be distorted when it is translated in words. To reduce the errors caused by a mixture of subjects whose statements could be incredible, we contrived an experimental plan as follows.

3.1 Procedure

First, we had to prepare a test piece, the grouping to which may differ according to which schema is predominant. To put it concretely, we used a musical piece whose grouping would differ according to whether the listener paid attention to the melody or to the inner voice. For this purpose, we used performances of the first eight measures of the second movement of Beethoven’s Piano Sonata "Pathetique," explained and realized by Professor Hiroshi Hoshina, a composer and conductor [9]. In the experiment subjects were asked to mark any of the grouping candidates and any of the accent notes in the melody that they felt, into the given score sheet (see Fig. 3), after listening to each of the following three stimulus performances.

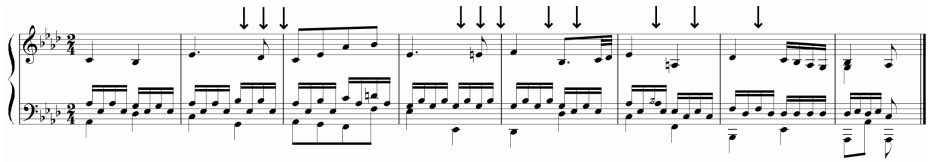


Fig. 3. Score Sheet for the Listening Experiment: Arrows are candidates of grouping boundaries. After listening to each performance, subjects mark their grouping boundaries and accented notes that they felt.

The first stimulus was an **original expressive version**: a performance including all dynamics and tempo. The second stimulus was a **chord version**: a performance whose accompaniment was replaced with chords synchronous to the melody. The player’s inner voice expression was thus suppressed while maintaining the harmonic structures of that piece. The third stimulus was the **original expressive version**, again. Before the third trial, subjects were given an explanation of the roles of the inner voice in the piece, that is, the part that gives the piece phrasing expression (see Fig. 4).

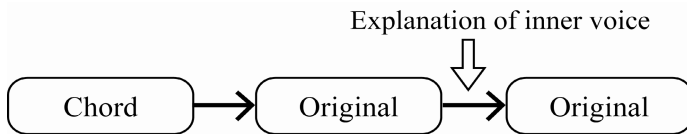


Fig. 4. Process of the Experiment

3.2 On the 2nd Movement of Beethoven’s “Pathetique”

The second movement of “Pathetique” is representative of homophony. Homophonic music has a main melody and a synchronized homogeneous accompaniment. According to Hoshina’s analysis, the melody of the first eight measures consists of six groups supported by the melodic and harmonic solution (brackets located at the upper part of the score in Fig. 5). The second and third groups (from C in the first beat of the third measure to F in the first beat of the fifth measure) combine to form a compound group. The root chord in the third group keeps V (F at the last note means V9), and

the first note (E flat) chromatically progresses to the last F (that is, the middle E natural is a passing note). Based on this analysis for music groups and considering the contour, Hoshina identified the apex note, marked by a star within each group. The fundamental expression of the groups that Hoshina suggests is of crescendo toward the apex note, and decrescendo from the apex to the end of the group, which is regarded as typical phrasing.

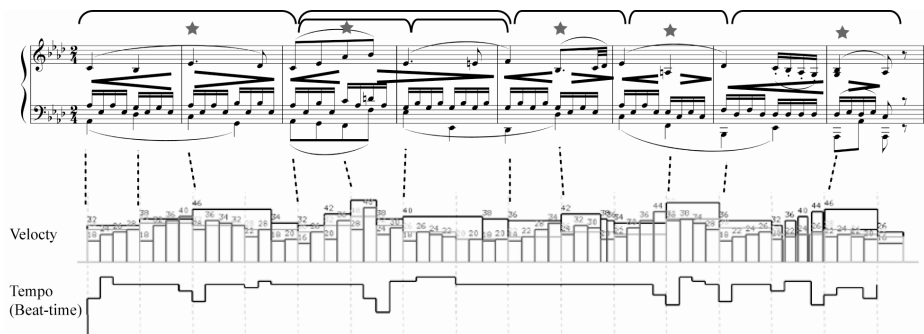


Fig. 5. Interpretation and Expression of 2nd Movement of Beethoven's Piano Sonata "Pathétique": The annotated slurs are according to the Henle Edition. The brackets above the main melody mean Hoshina's grouping and the stars mean the apices of those groups. Crescendo and diminuendo are written by the author based on his expression.

In this music, 16th notes of the inner voice, especially, 16th notes from the 1st measure to 2nd measure, the 4th measure, and the 6th measure, of which the melody notes are long tones, play an important role in phrasing expression. It is impossible to express gradual crescendo within a note played by a piano. Phrasing for group expression is realized with crescendo followed by diminuendo for the part described in section 2.1.2. Therefore, phrasing of these groups is achieved with a crescendo followed by a diminuendo for the 16th notes of the inner voice.

The 1st note of the 6th measure is a boundary candidate in the **AOG schema**. The melody note E flat is the last note of the preceding group and is the starting note of the next group (6th bar). It is impossible to express crescendo only with the expression of the summit note (A natural) in the melody. In contrast, in the preceding group, the summit note of the melody B flat (The 2nd beat in the 5th measure) and resolved to E flat with a diminuendo. It is possible to express diminuendo for the ascending 32nd note sequence. To sum up, A natural (2nd note of the 6th measure) is likely to be a group start because this note is adjacent to the preceding note sequence and the intensity of A natural is stronger than that of E flat.

The discussion so far can be summarized as follows; (a) if a listener finds that A natural (the 2nd note of the 6th measure) is the group starting note, he or she has used the **AOG schema**. (b) If the listener does not regard A natural as the group starting note, that is, s/he has grouped by **phrasing schema**, s/he must be tracing the inner voice, whether s/he may be conscious of that, or not. If the inner voice is replaced with chords, some of the (b) listeners will indicate a group starting at the A natural note.

Hoshina analyzed the performance interpretation of eight measures of the beginning of the second movement of this piano sonata, describing its dynamics [velocity] and tempo [BPM] by sixteenth note. His grouping, apices, and performance information are shown in Fig. 5. The information does not include pedal or note length. As for this piece, pedaling is not always required. In addition, this piece is often played slowly with the sounds of the piano sustained acoustically. Thus, we made each note length equal to its duration, except for notes annotated with a slur and staccato, which were given eighty percent of their own duration. We used a YAMAHA MU-2000 instead of the MU-50 used by Hoshina for the output of the two stimuli, and converted the output into MP3. Appendices 1 and 2 show the performance data.

3.3 Results

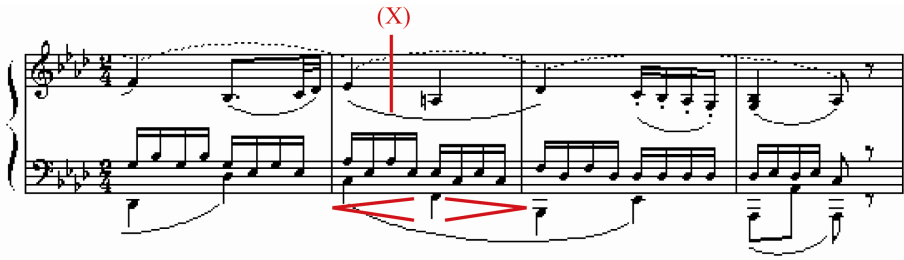
The subjects were 231 university students. The boundary that the subjects marked most, when they heard the first original expressive version, was between E flat and E in the 4th measure (59%). The second dominant boundary was just before the starting note in the 7th measure (15%) and the third dominant boundary was between the 1st beat and the 2nd beat in the 6th measure (9.5%).

The top and the second dominant boundaries the subjects chose are the position where a half cadence pattern appears and the position where a typical cadence pattern starts, respectively. On the other hand, it is not impossible to explain why the third position was chosen only with the cadence patterns.

Here we focus on the third dominant boundary, called X in the remainder of this paper, where the intense E flat may lead to an AOG schema and phrasing achieved with an expression of the inner voice may lead to a phrasing schema. Among 231 subjects, 20 subjects said there was a boundary at X after listening to the chord performance. Among these 20 subjects, 14 subjects said that they did not notice the boundary at X when they listened to the source expressive performance, while the other 6 said they did. The difference between the source expressive performance and the chord performance is whether the notes in the inner voice are played with the sequential 16th notes or played at the same time. 14 subjects (**FA_I** Group) were regarded as listeners possessing both **AOG schema** and **phrasing schema**, and as paying attention to the inner voice. They are supposed to be listeners of traditional western music.

Sixteen subjects thought there was no boundary after the inner voice function was explained. These listeners were deemed to be ones whose **AOG schema** is somewhat prior to their **phrasing schema**, and whose attention to the inner voice is comparatively low (**AF** Group).

Six subjects who said that they felt there was a boundary after being given the explanation were regarded as listeners with a strong **AOG schema** (**A** Group). If we make experiments using other musical samples, we might be able to separate the remaining 188 subjects' musical grouping types more precisely. However, it is impossible to judge whether the remaining subjects have perceived a group boundary and produced consistent statements, only from the experiment using Beethoven's Piano Sonata "Pathétique". Therefore, we are going to consider the subjects of the **FA_I**, **AF** and **A** group in the following discussion.



	Chord	Original	Original after explanation
subjects who feel a boundary at (X)	20	22	17
subjects who didn't feel a boundary at (X)	211	209	214

Transitions: 20 to 22 (14 FA_l), 22 to 17 (6 A), 211 to 209 (16 FA_l), 209 to 214 (16 AF)

Fig. 6. Result of Boundary Perception between the 1st beat and 2nd beat at the 6th bar

4 Grouping Schema and Its Background

The investigation of causes that lead to formulation of individual schema is a crucial study target of Kansei research. However, it is not easy to control individual conditions that might affect schema formulation. In the experiment, we investigated (a) preference for musical genre and (b) preference for listening manner, based on the social-psychological approach of Csikszentmihalyi [17].

The procedure we adopted is that subjects should give the order to items that we prepared in advance, as shown in Table 1. We then analyzed the preference characteristics of each group (FA_l, AF Group, A Group) by calculating the average preference order. The average order and the radar charts are shown in Table 2 and Fig. 7, respectively. The scale of the radar chart is given by 1.0 – average order / number of items. It is not an absolute scale. However it provides us with information to understand the tendency.

Music Genre Preference Result

- The preference level for Pop was the highest for all of the groups.
- The preference levels of the FA_l group were lower for Rock and Rap, compared with those of other groups, and higher for Therapy Music.
- The preference levels of the A group were lower for Bach and Chopin, compared with those of other groups.
- The preference levels of the AF group were between those of the FA_l group and A group.

Preference for Listening Manner

- Listeners of the FA_l group preferred listening music alone and quietly.
- Listeners of the A group preferred listening to music with their friends. They attached importance to lyrics.
- The preference levels of the AF group were between those of the FA_l group and A group.

Table 1. Questionnaire items of preference for music genre and listening manner

Favorites of music genres	
● Pop	● Healing (Therapy Music)
● Rock	● Typical Movie Soundtrack
● Rap	● Chopin
● Jazz	● Bach
● Dance	● Orchestra
● Ballad	● Other (Free Description)

Activity while listening	
● Listen in a lively place	● While concentrating
● Listen with friends	● While anticipating the progression of the music
● While riding the wave of Music	● Play the music in your mind
● Listen as BGM: background music	● While being emotionally affected
● Listen in a quiet place	● Listen to lyrics rather than melody
● Listen Alone	● Listen to melody rather than lyrics
	● While trying to remember the lyrics and melody

Table 2. Orders of preference for music genre and listening manner

Genre	FA _i	AF	A	Others	Listening Situation	FA _i	AF	A	Others
	(14)	(16)	(6)	(188)		(14)	(16)	(6)	(188)
Pop	2.64	2.19	2.00	2.45	Lively	9.79	8.94	6.50	9.34
Rock	4.54	3.94	2.83	4.66	With Friends	9.50	7.28	5.17	8.83
Rap	7.14	5.13	4.50	6.63	With Moving	7.50	6.28	8.00	8.13
Jazz	5.61	6.00	6.00	6.41	As Background Music	4.93	5.28	5.50	4.99
Dance	6.93	6.31	7.50	7.22	Quietly	3.57	5.97	7.50	4.21
Ballad	4.43	4.31	3.17	4.62	Alone	3.21	4.34	6.00	3.57
Healing	6.18	7.69	8.17	6.85	Concentrating	5.93	5.81	7.00	5.33
Typical Movie Soundtrack	6.11	6.06	4.67	5.43	Anticipating	7.57	7.34	9.17	7.88
Chopin	7.11	7.94	9.83	7.49	Imagine Playing	8.71	8.53	8.83	8.27
Bach	8.04	8.56	10.17	8.13	Emotionally Affected	9.57	9.06	9.33	9.64
Orchestra	8.25	9.00	8.50	7.76	Listen to Lyrics	8.71	8.31	6.67	8.00
Others	11.04	10.69	10.67	10.21	Listen to Melody	5.64	6.00	4.67	5.23
					Trying to remember	6.36	7.09	6.67	7.47

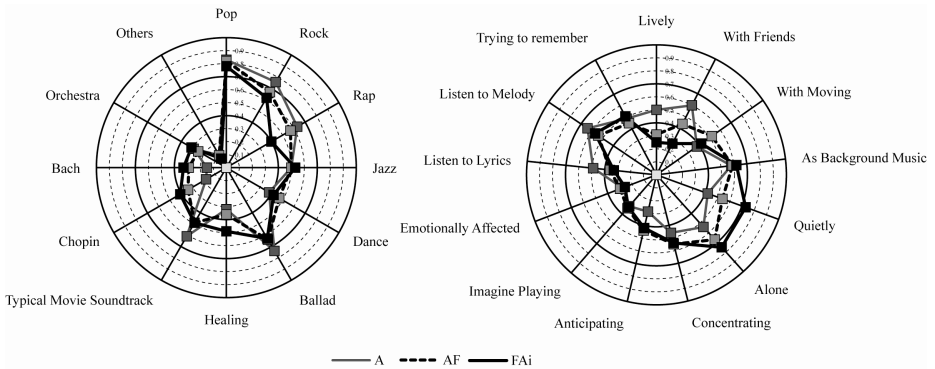


Fig. 7. Radar chart of Table 2. The preferred items are listed on the circumference.

FA₇ Group. The listeners of the FA₇ group are supposed to be interested in traditional western music, because they can hear the inner voice. We hypothesized that the genres preferred by this group would be orchestral music, Bach, and Chopin. Unexpectedly, the results show that the preference for orchestral music was not so high. However, the preferences for Bach and Chopin were higher and preferences for Rock and Jazz were lower than in the A group. Listeners of this group tended to listen to music alone and quietly. They did not prefer listening with many people. They seemed to enjoy listening only to the music itself. Despite the slight deviation from our expectation, we may say the hypothesis was supported by the results on the whole.

A Group. Listeners of the A group are thought to have the AOG schema. We hypothesized that they would have a strong preference for Rap and Rock music, in which the beat expression is prior to phrasing. Fig. 6. shows that the listeners of this group do prefer Rap and Rock music, and have little interest in Bach and Chopin. They prefer listening to music with their friends. In addition, the listeners of this group regarded lyrics as important, and they make light of anticipation and concentration, compared with other groups. We may reasonably conclude that the hypothesis was supported by the results.

5 Discussion

Our discussion of the results can be summarized as follows:

In the experiment, we focused on two primary grouping schemata, and gave higher priority to excluding subjects whose statements were inconsistent. Eventually, the 90% of the subjects that did not find a boundary at the first beat of 6th measure were taken out of the analysis. There is a possibility that other schema view this part as not the boundary, for instance, a temporal note line schema. We would like to conduct further experiments using composed music to investigate this possibility.

The questionnaire items were limited to coarse items except for Bach and Chopin. This is because we considered the subjects would find it more convenient when filling in the form if we selected the items to be evaluated. For the same reason, we adopted an ordinal scale instead of an interval scale. We should thus conduct future investigations based on an interval scale, limiting the number of items.

The questionnaire survey revealed that musical preferences are different if the listener type (i.e. **A, AF, FA₇ group**) is different. At the same time, we found preference differences corresponding to age. That means we should conduct broader investigations with a greater range of subject ages.

6 Conclusion

The music grouping depended on each listener's musical experience. We provided a working hypothesis that grouping differences result from the balance of schemata strengths that each listener possesses, and executed an experiment using the second movement of the "Pathétique" composed by Beethoven. The experimental parts of music listened to by the subjects were controlled, as the grouping results might differ according to which of the **accent-oriented schema** and **phrasing schema** is prior for

the subject. We were able to identify three typical listener types; those who have strong **accent-oriented schema** (**A** group), those whose **accent-oriented schema** is somewhat prior to their phrasing schema (**AF** group), and those who have the competence of listening to inner voices and whose **phrasing schema** is prior to their accent-oriented schema (**FA_I** group).

Using a questionnaire survey filled out by these subjects, we verified that (1) **A** group listeners prefer Rap music and Rock music, as well as listening in a lively place with many people and listening to lyrics. (2) **FA_I** group listeners prefer Chopin, listening alone and quietly, and (3) university freshmen in the **FA_I** group do not have as much preference for Rap or Rock compared with freshmen in other groups.

The subject dealt with in this paper is an investigation of human cognitive processes based on observation of the musical grouping process. We illustrated that we can classify the listener's type considering superiority of grouping schemata. We also investigated the relationship between the listeners' schema type and their musical preference. Our investigations are still at an early stage, and we would like to conduct broad investigations together with experiments using samples of originally composed music as future work.

References

1. Koffka, K.: Principles of Gestalt Psychology, Harcourt (1967)
2. Ozawa, S., et al.: Lecture notes of Hideo Saito, Hakusui-sha (in Japanese, 1999)
3. Hiraga, R., Hashida, M., Hirata, K., Katayose, H., Noike, K.: RENCON: toward a new evaluation method for performance rendering systems, Proc. Int. Computer Music Conference (2002) 357-360
4. Noike, K., Hashida, M., Katayose, H.: An Examination Tools to obtain Musical Group Boundary --- WebMorton, IPSJ 2002-MUS-49 (in Japanese, 2002) 25-29
5. Katayose, H., Hashida, M., Noike, K.: On Perception of Apices and Group Boundaries in Listening Music, IPSJ 2003-MUS-52 (in Japanese, 2003) 95-102
6. Lerdahl, F., Jackendoff, R.: A Generative Theory of Tonal Music, MIT Press (1983)
7. Murao, T.: Recognition of Music Analysis, Music and Recognition, University of Tokyo Press (in Japanese, 1987) 1-40
8. Aiello, R.: Musical Perceptions, Oxford University Press (1994)
9. Hoshina, H.: Approach to lively musical expression, Ongaku-no-tomo-sha Corp., (in Japanese, 1998)
10. Snyder, B.: Music and Memory: An Introduction, MIT Press (2001)
11. Moore, B. C. J.: An Introduction to the Psychology of Hearing, Academic Press (2003)
12. Edelman, G. M., Gall, W. E.: Auditory Function, Neurobiological Bases of Hearing, John Wiley & Sons Inc (1988)
13. Hamanaka, M., Hirata, K., Tojo, S.: Automatic Generation of Grouping Structure based on the GTTM, Proc. Int. Computer Music Conference (2004) 141-144
14. Noike, K., Hashida, M., Takeuchi, Y., Katayose, H.: Expansion of GTTM Grouping Rules for Involving Expression Parameters, 2004-MUS-57 (in Japanese, 2004) 11-16
15. Takeuchi, Y.: Performance Variables for Grouping Structure in two editions of the theme of K. 331, Rencon Workshop, Proc. of ICAD (2002)
16. The New GROVE Dictionary of Music and Musicians, Groves Dictionaries Inc. (1993)
17. Csikszentmihalyi, M.: Beyond Boredom and Anxiety, Jossey-Bass Inc. Pub. (2000)