Audio Engineering Society

Convention Paper 9950

Presented at the 144th Convention
2018 May 23 – 26, Milan, Italy

Evaluation of Player-controlled Flute Timbre by Flute Players and Non-Flute Players

Mayu Kasahara[1], Atsushi Marui[2], and Toru Kamekawa[2]

[1]*Graduate School of Music, Tokyo University of the Arts*
[2]*Faculty of Music, Tokyo University of the Arts*

Correspondence should be addressed to Mayu Kasahara (`mayucopolo1219@gmail.com`)

ABSTRACT

In order to investigate how flute players and non-flute players differ in the perception of the instrument, two listening experiments were carried out. The flute sounds were recorded to have changes in five levels of harmonic overtones energy levels played by three flute players. Through listening experiment of attribute rating on "brightness," the flute players were found to evaluate the stimuli "brighter" as the harmonic overtones energy decreased while the non-flute players evaluated inversely. Through the second listening experiment of pairwise global dissimilarity rating among the stimuli, two dimensions corresponding to the harmonic overtones energy levels and to the noise levels were found. The experience of the flute performance did not seem to affect the result. These results indicate that the experience of the flute performance seemed to affect the result only when evaluating the stimuli using the word "brightness."

1 Introduction

Flute players are capable of controlling the number and level of harmonic overtones, when playing the lower range of the instrument, by changing the angle of the instrument, the velocity of the air beam, and the embouchure (mouth shape). Thus, the total energy of the harmonics over the fundamental changes, resulting in a difference of timbre. In this paper, it is referred to as "harmonics energy level." Specifically, when the angle of the instrument is tilted inward or the velocity of the air beam is increased, the harmonics energy level increases and when tilting the instrument to the outside or breathing upward (*i.e.*, outside the pipe), the harmonics energy level decreases.

Most of the studies of the flute have focused on physical aspects such as the tone generating mechanism of air-reed. Research on the edge tone, the sound generated from the jet colliding with the edge was begun in the first half of the 20th century, and research on the sounding mechanism of the air-reed instruments was started in 1960–70's. For example, Coltman [1] proposed the theory about the oscillation of the air-reed instruments. Fletcher and Rossing [2] developed the theory about the behavior of air jet and the mechanism of resonance in detail based on the study of Coltman *et al.*. Also, Ando [3, 4, 5] investigated the relationship of the drive condition of the flute and the generated sound and proposed that the bias of the air beam (how much the center plane of the thickness of the air beam leaving the lip is biased toward the inside of the pipe) and the velocity of the air beam are considered to be the most important conditions among the nine drive conditions such as the velocity and the thickness of the air beam, the radius of the upper lip, *etc.*, to control the physical parameters of the generated sounds. In these studies, Ando confirmed that the velocity of the air beam affects the sound pressure of the generated sound in the whole range and the abundance of the harmonic overtones in the lower range, and that it plays an important role that harmonic overtones become abundant with the velocity of the air beam in the lower range to control the loudness in this range.

Grey [6] investigated the timbre of musical instruments from a more subjective viewpoint. He examined the perceptual similarities of 16 musical instrument samples including a flute tone that were synthesized based upon an analysis of actual instrument tones, and the result was analyzed with multidimensional scaling and hierarchic clustering. A three-dimensional timbre space was obtained and each dimension was corresponded to 1) the spectral energy distribution; 2) the presence of synchronicity in the transients of the higher harmonics, along with the closely related amount of spectral fluctuation within the the tone through time; and 3) the presence of low-amplitude high-frequency energy in the initial attack segment. In his study, the flute sound was classified into the same cluster as the three cello sounds played in different position of bow.

Although the numerous studies on the flute timbres exist, authors know of few studies referred to the player-controlled harmonics energy level such as the above-mentioned works by Ando [3, 4, 5]. Investigation of the harmonics energy level and its relation to the timbre perception are not thoroughly done, and there are few reports about it other than the personal opinions of the flute players or the listeners. However, their opinions do not always agree. For example, non-flute players usually evaluate the flute timbre as brighter along with the increase of the harmonics energy level. This corresponds with the former study suggesting brighter evaluation of sounds is generally caused by the higher spectral centroids and the existence of higher harmonics [7, 8, 9]. In contrast, flute players tend to evaluate the timbre brighter as the harmonics energy level decreases as far as the first author inquired about. In lessons and recordings, the timbre is often expressed in words, but if there is a difference in perception of the timbre between players and non-players, it may become an obstacle for communication between a performer and a recording engineer. Investigating how people feel the player-controlled flute timbres by a quantitative method may contribute to smooth communication in these situations.

In this study, two listening experiments were carried out for the purpose of revealing how flute players and non-flute players perceive the player-controlled flute timbres.

Section 2 of this paper describes sound stimuli whose harmonics energy levels were controlled by the flute players. Section 3 presents about experiment 1 examined the brightness of the stimuli, section 4 presents about experiment 2 evaluated global dissimilarity among the stimuli, and section 5 discusses about these results.

2 Sound Stimuli

Three students majoring in the flute in Tokyo University of the Arts or Tokyo College of Music participated to record the stimuli. The flutists were informally interviewed before the recording and responded that they can play tones in different levels of harmonic energy. Thus, they were asked to play long tones of D4 (295 Hz) and H4 (496 Hz) five times without vibrato and to increase the harmonics energy level gradually from the first time to the fifth time.

The players had around ten years of flute playing experiences. Player A played the Brannen, Player B played the Powell, and player C played body and foot-joint of the Muramatsu and head-joint of the Nagahara.

Total of 15 stimuli for D and H respectively were prepared. The loudnesses of the stimuli were equalized using the Replay Gain function of the Audacity audio editing software followed by the first author's fine adjustments. The stimuli were labeled as combinations of a player symbol (A, B, or C) and the level of five harmonics energy level (1 through 5 in increasing order). Spectral centroids of them are shown in Table 1. The spectral centroid was calculated using the following equation,

$$\text{Spec. Cent.} = \exp\left(\frac{\sum_{n=1}^{N} x_n \log f_n}{\sum_{n=1}^{N} x_n}\right) \tag{1}$$

Table 1: Stimuli codes and their spectral centroids in Hz (upper panel: D tones, lower panel: H tones)

Player A	A1	A2	A3	A4	A5
	303	373	391	454	546
Player B	B1	B2	B3	B4	B5
	345	437	452	496	627
Player C	C1	C2	C3	C4	C5
	373	518	540	573	749
Player A	A1	A2	A3	A4	A5
	745	670	985	944	986
Player B	B1	B2	B3	B4	B5
	591	699	819	975	1097
Player C	C1	C2	C3	C4	C5
	545	568	744	907	981

where n is FFT bin number from N bins and x_n is an energy at bin frequency f_n of the power spectrum of a stimulus. Logarithm and exponentiation are used to roughly correspond the result to the perceived pitch of a frequency (note that multiplication in frequency scale corresponds to addition in the octave scale). Spectral centroids of the stimuli appeared to be increasing approximately corresponding to the order in which the players played.

3 Experiment 1

3.1 Method

Scheffe's pairwise comparison method using seven-point scale was used. Participants were asked to evaluate which and by how much of the two presented stimuli is brighter. The term *akarusa* ("brightness" in Japanese) was used because all participants were native speakers of Japanese language. The total number of pairwise comparisons was 105 ($= 15 \times (15-1) \div 2$) each for D and H tones.

The GUI for the experiment (Fig. 1) was programmed with Pure Data. The participant was allowed to switch between the two stimuli at any time using the two boxes with a filled box (under "Switcher"), and responded with the seven-point scale under it. Stimuli were looped until the participant pressed a button to proceed to the next pair of stimuli.

The stimuli were presented via a pair of Sony MDR-CD900ST headphones connected to the analogue audio output jack of an Apple MacBook Pro (2013 model) placed in a quiet room. Before the experiment, participants listened to all stimuli in a random order to make a basis for evaluation. Loudness level was adjusted by each participant before the experiment and was not changed during the experiment.

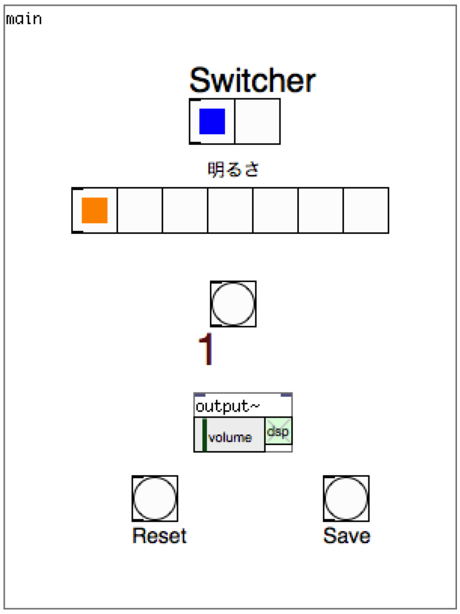

Fig. 1: Graphical User Interface used in the Experiment 1. The Japanese characters above the seven-point selector reads *akarusa* ("brightness").

3.2 Participants

Students of the Tokyo University of the Arts participated in the experiments. They were divided into three groups according to their musical experiences. The first group with seven flute players who have nine to eleven years of playing experience was named "Flutist group." The second group with twenty non-flute players and eighteen of those were majoring in a sound recording technology and acoustics research department was named "Engineer group." The third group with seven non-flute players majoring other instruments was named "Musician group." Three of them were majoring piano, one was majoring contrabass, one was majoring horn, and the other two were majoring *shakuhachi* (Japanese end-blown flute).

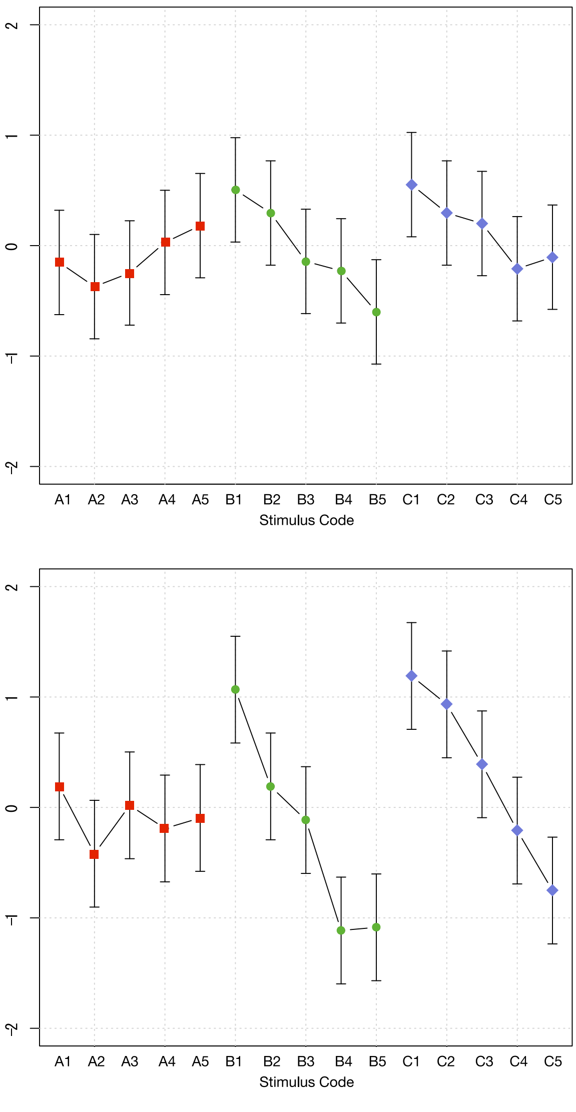

Fig. 2: Ratings by the Flutist group (upper panel: D tones, lower panel: H tones). Symbols show mean ratings (higher is "brighter") and error bars show 95% confidence intervals.

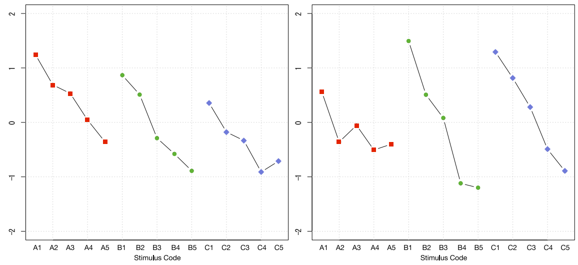

Fig. 3: Mean ratings of the Type 1 of the Flutist group (left panel: D tones, right panel: H tones)

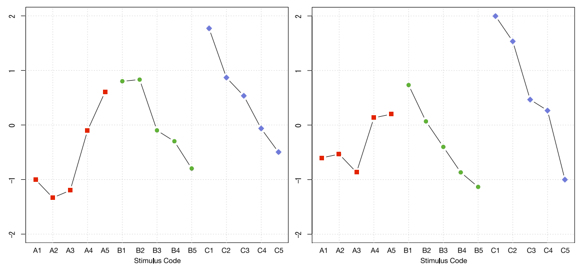

Fig. 4: Mean ratings of the Type 2 of the Flutist group (left panel: D tones, right panel: H tones)

3.3 Result

The result of the Flutist group is shown in Fig. 2. The horizontal axis represents the stimulus number. The vertical axis represents the rating of brightness, and it is evaluated as brighter as the stimulus is located above.

The Flutist group indicated the tendency of evaluating the stimuli "brighter" as the harmonics energy level decreased. However, the graph of Player A appears to the ascending in D tones and the slope is not as steep as the other players in H tones. It is possible that the mean ratings were offset because of individual differences in the evaluation of seven participants. Therefore, seven people were classified into three types for each tendency of individual evaluation.

Type 1 is the type that evaluated the stimuli "brighter" as the harmonics energy level decreased (Fig. 3). Three people in D tones and five in H tones came under this type. Type 2 is those who evaluated the stimuli of player A "brighter" as the harmonics energy levels increased, and the stimuli of the other players "brighter" as the harmonics energy level decreased (Fig. 4). Two people in D tones and one in H tones came under this type. Type 3 is those who evaluated the stimuli

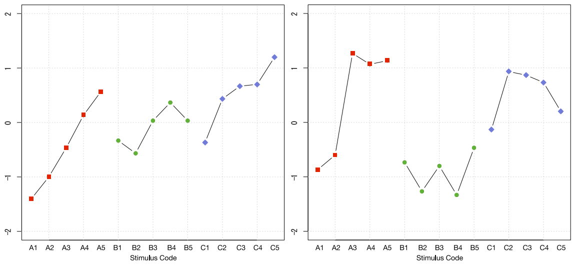

Fig. 5: Mean ratings of the Type 3 of the Flutist group (left panel: D tones, right panel: H tones)

"brighter" as the harmonics energy levels increased (Fig. 5). Two people in D tones and one in H tones came under this type.

Since non-flute players of the Engineer group and the Musician group showed the same tendency of evaluation, the result unified both groups is shown in Fig. 6. The Engineer and Musician groups indicated the tendency of evaluating the stimuli "brighter" as the harmonics energy levels increased in contrast to the overall trend of the Flutist group.

From these results, it was found that the flute players evaluate the stimuli "brighter" as the harmonics energy level decreases, although there were individual differences, and the non-flute players evaluate the stimuli "brighter" as the harmonics energy level increases.

4 Experiment 2

4.1 Method

Experiment 2 was conducted using same method and the stimuli as experiment 1. Participants were asked to evaluate the pairwise global dissimilarities of stimuli's timbre in seven-point dissimilarity scale. The reason behind using global timbral differences is that to avoid the use of verbal attributes.

4.2 Participants and Environment

Five participants from the Flutist group and fourteen participants from the Engineer group participated in the experiment 2. Two new participants joined in the Engineer group for the experiment 2. Environment and the equipment were the same as in experiment 1.

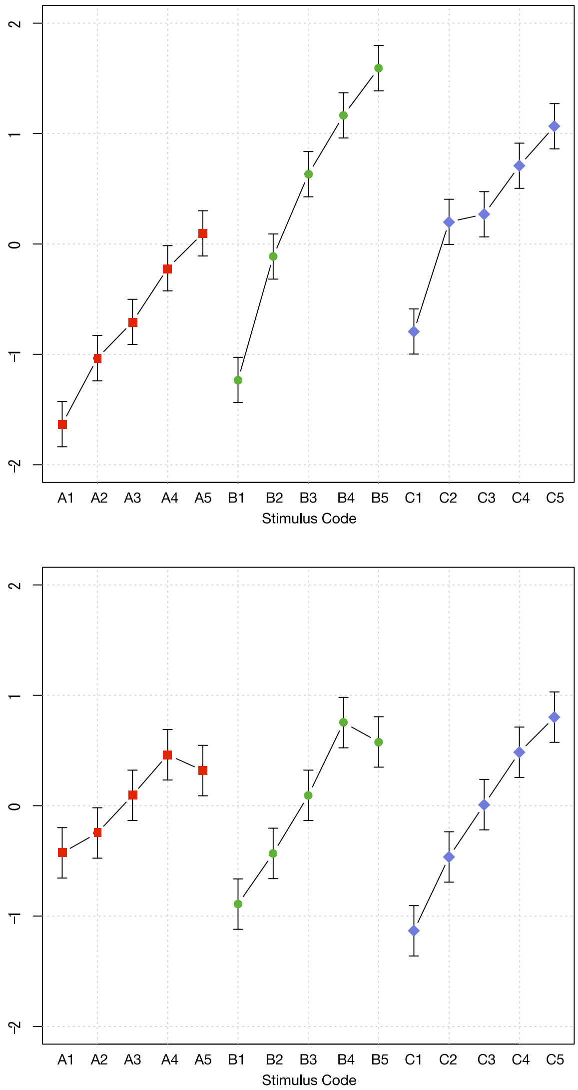

Fig. 6: Ratings by the Engineer group and the Musician group pooled (upper panel: D tones, lower panel: H tones). Symbols show mean ratings (higher is "brighter") and error bars show 95% confidence intervals.

Table 2: Stimuli codes and their Harmonics-to-Fundamental Energy Ratio in dB (upper panel: D tones, lower panel: H tones)

Player A	A1	A2	A3	A4	A5
	2.36	8.40	10.40	12.59	15.23
Player B	B1	B2	B3	B4	B5
	6.97	12.16	13.17	14.39	15.23
Player C	C1	C2	C3	C4	C5
	6.74	14.30	15.08	16.22	17.31
Player A	A1	A2	A3	A4	A5
	9.49	8.15	13.29	13.06	14.54
Player B	B1	B2	B3	B4	B5
	3.88	8.86	10.43	13.28	14.91
Player C	C1	C2	C3	C4	C5
	1.71	4.44	9.99	14.01	15.17

4.3 Result

First, collected global differences data were submitted to Hierarchical Cluster Analysis to see whether participants' individual differences were related to the group they belong to. From the result, it was suggested that the effect of group differences was small, thus the obtained global differences data were pooled and analyzed with Multidimensional Scaling Analysis.

The result of the analysis of the MDS revealed two dimensions in the resulting stimulus space as shown in Fig. 7. Stimuli of different harmonic energy levels are generally ordered along the first dimension. Player differences can be seen along the second dimension.

Measures derived from the stimuli were examined. For the first dimension, each stimulus was divided into low- and high-frequency bands at the frequency between the fundamental and the second harmonic (350 Hz for D tones and 750 Hz for H tones). Then the measure indicating the energy ratio of the high-frequency band and the low-frequency band in dB was calculated (Table 2). Although the stimuli include inharmonic components, it represents the approximate energy ratio of the harmonics and the fundamental of each stimulus. We refer this index as HFR (Harmonic-to-Fundamental energy Ratio). HFR becomes higher with the greater energy of the high-frequency bands relative to the energy below and including the fundamental frequency. HFRs correlated well with the first dimension ($r = .94$ and $r = .95$ for D and H tones respectively).

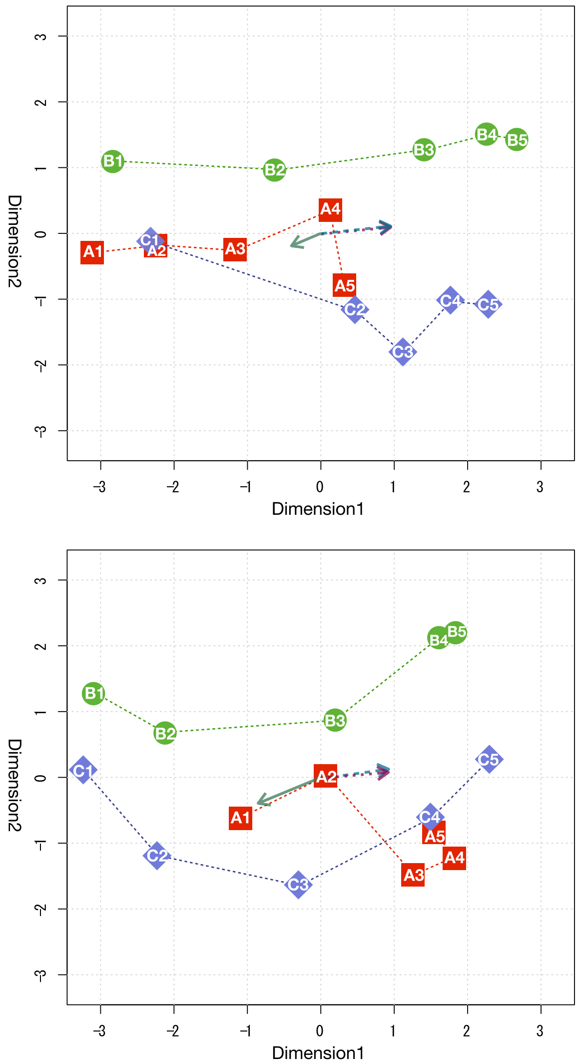

Fig. 7: Two dimensional MDS solution of experiment 2 (upper panel: D tones, lower panel: H tones). The correlation coefficients between the "brightness" rating and each dimensions of MDS are shown as arrows (solid: Flutist group, dashed: Engineer group, dotted: Musician group).

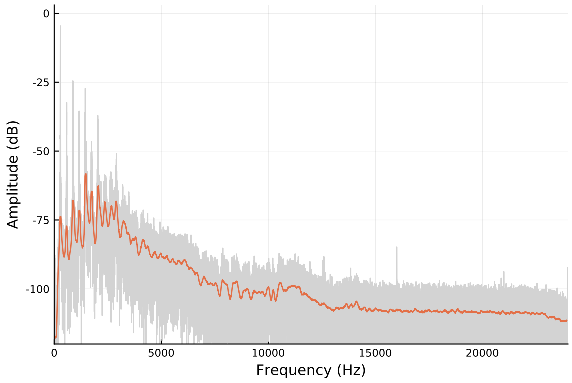

Fig. 8: Noise level calculation using moving median on the spectrum. Light colored lines are the power spectrum of the sound A1, and darker curve is moving median of the spectrum.

For the second dimension, moving median values of the spectrum was calculated. This was done by taking a part of spectrum that has the bandwidth twice the fundamental frequency and the median value of the power was calculated, which was repeated by sliding the band through the spectrum. An example result for stimulus A1 is shown in Fig. 8. Light colored lines are the power spectrum and darker curve is the moving median of the spectrum. It can be seen that the peaks of the spectrum (*i.e.*, harmonic overtones) are smoothed out while retaining the inharmonic parts. The sum of this moving median can be seen as the approximation of the inharmonic energy levels which is related to "breathiness" of a flute tone. We call this index "noise level," and the calculated values for the stimuli are shown in Table 3. The second dimension correlated well with these inharmonic noise levels ($r = .89$ and $r = .73$ for D and H tones respectively).

Through the above analyses, it was found that the harmonics energy level corresponds to the dimension 1 and the noise level corresponds to the dimension 2.

5 Discussion

The correlation coefficients between the brightness ratings in experiment 1 and MDS dimensions obtained in experiment 2 was calculated. Fig. 7 shows the correlation between the rating of brightness of each group and each dimension on the MDS solution as arrows (solid-lines: Flutist group, dashed: Engineer group, dotted: Musician group). The coordinate values of the tips of the arrows correspond to the correlation coefficients. The dimension 1 had a relatively high correlation with brightness ratings in any group and dimension 2 had low correlation coefficient.

Table 3: Stimuli codes and their Noise Levels in dB (upper panel: D tones, lower panel: H tones)

Player A	A1	A2	A3	A4	A5
	61.55	59.97	60.53	60.07	60.87
Player B	B1	B2	B3	B4	B5
	61.73	63.34	62.96	64.14	66.47
Player C	C1	C2	C3	C4	C5
	59.66	58.97	57.37	57.24	57.46
Player A	A1	A2	A3	A4	A5
	55.96	53.66	52.58	53.92	54.74
Player B	B1	B2	B3	B4	B5
	56.11	58.94	57.85	57.97	58.04
Player C	C1	C2	C3	C4	C5
	53.70	53.58	55.98	54.36	56.47

However, if we see the correlation values for individual participants, ratings of some participants had high correlation value with dimension 2. In H tones, the maximum absolute value of correlation coefficient between rating and dimension 2 was $r = .79$. Therefore, there seems to some participants exist focused on the noise level when evaluating the brightness.

Through the two experiments, the experiences of the flute performance affected the result only when evaluating the stimuli using the word "brightness." However, the reason why the evaluation on the brightness by the flute player is different from non-performers is yet to be understood. There may be two hypotheses for explaining this results. First is that the participants listened to the same parts of the stimuli regardless of whether he/she is a flute player or not, but the flute players evaluated it in conjunction with the feeling when they play the instrument. When playing the flute, if they want to suppress the harmonics energy levels, they loosen the embouchure or breathe slightly upwards. This may be related to the word of positive impression (in this case "bright"), and thus the word may be used in the different meaning from the other groups. Second is that the flute players were evaluating based on the different acoustic features from non-flute players. In that case, it is necessary to investigate the physical features corresponding to the perception of the brightness for the flute player. However, further investigation is necessary to understand the reason why the evaluation on the brightness by the flute player is different from non-performers.

6 Conclusion

In order to investigate how flute players and non-flute players differ in the perception of the instrument, two listening experiments were carried out.

In the first experiment, participants were asked to compare the flute sound stimuli with five levels of player-controlled harmonics energy levels. Non-flute players evaluated the stimuli with more harmonics as brighter. Flute players tended to evaluate stimuli with less harmonics as brighter. However, some participants in Flutist group evaluated the stimuli with more harmonics bright, and others evaluated the stimuli with more harmonics played by a specific performer as brighter.

In the second experiment, the two-dimensional stimulus space was obtained from dissimilarity rating results. Dimension 1 corresponds to the harmonics energy level, and Dimension 2 corresponds to the level of the noise component included in each stimulus. Unlike Experiment 1, no influence of the flute performance experience of the participants was observed in this result.

It was indicated from the correlation coefficients between the brightness ratings in experiment 1 and MDS dimensions obtained in experiment 2 that most participants focused on the HFR when evaluating the brightness while some participants focused on the noise level.

There may be two hypotheses for explaining the reason why the flute players evaluated differently from the non-flute players when evaluating the stimuli using the word “brightness.” First is that the word may be used in the different meaning from the other groups, and second is that the flute players were evaluating based on the different acoustic features from non-flute players.

Although the further investigation will be needed for the difference in the evaluation between flute players and non-performers, this study may contribute to the communication between them in situations such as recordings.

References

[1] Coltman, J. W., “Sounding mechanism of the flute and organ pipe,” *Journal of the Acoustical Society of America*, 44(4), pp. 983–992, 1968.

[2] Fletcher, N. H. and Rossing, T. D., *The Physics of Musical Instruments*, Springer, 2nd edition, 1998.

[3] Ando, Y., “Drive Conditions of a Flute and their Influences upon Sound Pressure Level and Fundamental Frequency of Generated Tone (An Experimental Study of a Flute I),” *Journal of the Acoustical Society of Japan*, 26(6), pp. 253–260, 1969.

[4] Ando, Y., “Drive conditions of a flute and their influences upon harmonic structure of generated tone (An experimental study of a Flute II),” *Journal of the Acoustical Society of Japan*, 26(7), pp. 297–305, 1970.

[5] Ando, Y., *Acoustics of Musical Instruments, New Edition*, Ongaku no Tomo Sha, 1996.

[6] Grey, J. M., “Multidimensional perceptual scaling of musical timbres,” *Journal of the Acoustical Society of America*, 61(5), pp. 1207–1277, 1977.

[7] von Bismarck, G., “Timbre of Steady Sounds: A Factorial Investigation of Its Verbal Attributes,” *Acoustica*, 30, pp. 146–159, 1974.

[8] von Bismarck, G., “Sharpness as an Attribute of the Timbre of Steady Sounds,” *Acoustica*, 30, pp. 159–172, 1974.

[9] Yamada, M. and Miura, M., “Possibility of audio parameters as descriptor of musical audio,” *Journal of the Acoustical Society of Japan*, 70(8), pp. 440–445, 2014.