Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Study on effect of room acoustics on timbral brightness of clarinet tones. Part I: subjective evaluation through a listening experiment

Kosuke Kato (1), Tsubasa Nagao (2), Toshio Yamanaka (3), Keiji Kawai (4) and Ken-Ichi Sakakibara (5)

(1) Center for Advanced Science and Innovation, Osaka University, Japan
(2) School of Engineering, Osaka University, Japan
(3) Graduate School of Engineering, Osaka University, Japan
(4) Graduate School of Science and Technology, Kumamoto University, Japan
(5) School of Psychological Science, Health Sciences University of Hokkaido, Japan

PACS: 43.55.Hy, 43.66.Jh, 43.75.Pq

ABSTRACT

Musicians and acoustic engineers are often interested in knowing how their tone quality is related to the acoustics of a concert hall. This paper reports a listening test performed to investigate whether the effect of room acoustic conditions on the timbre of musical sound differs with the performing style of the produced musical sound. For semi-anechoic stimuli, nine natural clarinet tones produced at three different dynamic levels and three different notes (A3 ≈ 220 Hz, A4 ≈ 440 Hz, and A5 ≈ 880 Hz) were extracted from the RWC music database, which stores musical instrument sounds. For reverberant stimuli, 18 tones were generated by convolving each semi-anechoic tone with two different binaural impulse responses; these were collected at different seat positions in two different medium-sized concert halls. Fifteen instrumentalists, all with at least six years of experience in playing their respective instruments, participated in the listening experiment. The stimuli were presented dichotically over Sennheiser HD650 headphones to the participants at approximately 60 dB SPL for equal loudness in a soundproof room. The scale values of timbral brightness for the respective semi-anechoic stimuli and their reverberant stimuli at equal notes were obtained through Scheffe's paired-comparison test. Three sessions were carried out, and each session consisted of stimuli with the same tone. The results showed that the room acoustic conditions significantly affect the brightness of the clarinet tones, and the effect differs with the produced dynamic level and with the produced note. These findings call for a further examination of the interaction between the performing style and the room acoustics with respect to the brightness of clarinet tones; this will be presented in Part II of this study that is based on an acoustic analysis.

INTRODUCTION

In a concert hall, the audience listens to the musical sounds convolved with the room reverberation, while musicians adjust their performance to suit the acoustics for conveying their interpretation to the audience (Meyer, 2009; Ueno *et al*., 2010). Hence, the perception of sound in rooms may be closely related to both the room acoustic conditions and the acoustic characteristics of musical sounds. However, most studies on concert hall acoustics have conventionally assumed that the subjective effects in rooms can be described independent of the acoustic characteristics of the produced musical sound signals (e.g., Bech, 1995; Bech, 1996; Kuttruff, 2000; Zahorik, 2009).

If the effect of the room acoustic conditions (*room*) on the tone quality of musical sound differs with the acoustic characteristics of musical sounds, the subjective effects in rooms must be described in relation to the acoustic characteristics of musical sound. However, only a few attempts have been made to examine the timbre of musical sounds as perceived in a room in relation to the acoustic characteristics of musical sound signals (Hotehama *et al*., 2002; Disley and Howard, 2004; Hikichi and Miyoshi, 2006). To describe the timbre of musical sound as perceived in rooms, a systematic investigation of the interaction between the performing style and the room acoustics and the effect of room conditions on the timbre of musical sound is required.

The aim of this study was to investigate whether the *room* significantly affects the timbre of musical sound perceived in rooms and to investigate whether this perceived effect differs with the performing style of the produced musical sound.

MATERIALS AND METHOD

Choice of timbre adjectives commonly used by potential participants for listening experiment

In order to subjectively evaluate the timbre of musical sounds using adjectives, several timbre adjectives should be selected in advance of the listening experiment. To achieve this, two different approaches can be used. One is that the experimenter chooses and determines the timbre adjectives on the basis of his/her own experience and/or the results of previous studies. The other is that each participant is asked to describe and/or choose timbre adjectives from his/her own perspective (Kawai *et al*., 2004). The former is better at maintaining the generality of the experimental results, while the latter is better at avoiding ambiguous and/or unused adjectives for some of the participants. To maintain the generality of the experimental results and include words commonly used by all the participants, timbre adjectives commonly used among the potential participants for the following listening experiment were examined.

Forty Japanese-speaking instrumentalists—16 string instrumentalists, 13 wood wind instrumentalists, and 11 brass instrumentalists—performing as members of the same amateur orchestra group at Osaka University took part in a questionnaire survey.

A list of 84 adjectives, including 40 pairs of adjectives used in previous studies (Nanba and Kuwano, 1998) and 4 adjectives chosen based on one of the experimenter's experience was presented to each participant. Each participant was asked to judge his/her frequency of use for each adjective according to four levels: *yoku tsukau* (frequently used), *tokidoki tsukau* (sometimes used), *dochira tomo ienai* (neutral), and *tsukawanai* (unused).

The results showed that although the frequency of use for each adjective greatly depends on the instrumental group of participants, 90% of the participants for each instrumental group evaluated 10 of the frequently or sometimes used timbre adjectives: *akarui* (bright), *omoi* (heavy), *kitanai* (dirty), *komotta* (stuffy), *shin no aru* (core existing), *chikarazuyoi* (mighty), *hakkiritoshita* (distinct), *hibiki no aru* (reverberate or resonant), *fukami no aru* (deep), and *yawarakai* (soft).

Note that the perception of brightness has been shown to be simply described in relation to the spectral centroid of the musical sound signal (Schubert and Wolfe, 2006; Marozeau *et al*., 2003; Marozeau and Cheveigé, 2007). Thus, even in a room acoustic environment, "bright" may be not only a timbre adjective commonly used by the participants but also a perceptual dimension that is potentially described in relation to the spectral centroid. Therefore, this study began by investigating the perception of brightness in rooms.

Semi-anechoic stimuli

For semi-anechoic stimuli, nine natural clarinet tones produced at three different dynamic levels (*F*, *M*, and *P*), as shown in Table 1, and at three different notes (A3 ≈ 220 Hz, A4 ≈ 440 Hz, and A5 ≈ 880 Hz) were extracted from the RWC music database, which stores musical instrument sounds (Goto *et al*., 2003). The duration of each stimulus was between 2.1 and 3.1 s, as shown in Table 2.

Clarinet sounds produced at different dynamic levels were chosen as the stimuli for two reasons. One was that the clarinet's tone is normally produced without vibrato. Meyer (1993) reported that room acoustics affect the perception of the vibrato sound. It may be simpler to describe the perceived timbre in a room of musical tones without vibrato than with vibrato. The other reason was that the clarinet's tone may have one of the largest brightness ranges among instrumental sounds. Meyer (2009) reported that the produced dynamic level affects the timbral brightness of musical tone. He also indicated that the clarinet has one of the largest dynamic ranges among musical instruments. In this experiment, three different dynamic levels were chosen to explore the effect of the difference in dynamic level on the brightness perception of the clarinet sound in rooms.

Marozeau *et al*. (2003) and Marozeau and Cheveigé (2007) showed that the timbre of musical sound differs with the fundamental frequency (F_0). In order to examine the effect of the difference in F_0, clarinet sounds produced for three different notes were chosen.

Table 1. Produced dynamic levels and abbreviations

Dynamic level	Abbreviation
Piano	*P*
Mezzo	*M*
Forte	*F*

Table 2. Duration of each semi-anechoic stimulus

Note	Dynamic level	Duration [s]
A3	*P*	2.5
	M	3.1
	F	2.4
A4	*P*	2.4
	M	2.8
	F	2.1
A5	*P*	2.2
	M	2.6
	F	2.4

Reverberant stimuli

For reverberant stimuli, 18 tones were generated by convolving each semi-anechoic tone with two different binaural impulse responses (*BRIR*s), which were collected at a different seat position in two different medium-sized concert halls, as shown in Table 3. Table 4 and Figure 1 list the acoustic measures for the acoustic conditions of each room.

The two *BRIR*s were chosen based on a preliminary listening experiment done by one of the experimenters as follows. First, 22 *BRIR*s collected at 22 representative seat positions in the abovementioned two concert halls were convolved with each of the nine semi-anechoic stimuli. The spatial distribution for acoustical parameters in these two concert halls was previously analyzed in detail by a group including one of the current authors (Fujii *et al*., 2004). Second, the equivalent continuous A-weighted listening level for the sound-produced section (excluding the section for the reverberation tail) of each stimulus was adjusted to the same level through a signal processing technique. Third, the experimenter wore HD650 headphones connected to the headphone jack of a VAIO VGN-FJ11/W notebook computer. The experimenter then adjusted the sound reproduction level to her comfort level. The experimenter then randomly reproduced and compared each pair of semi-anechoic and reverberant stimuli. Finally, the experimenter chose two *BRIR*s that were perceived to have the largest and second largest effects on the brightness perception of musical sound among the 22 *BRIR*s.

Table 3. Abbreviation of the room acoustic conditions

Room acoustic condition	Abbreviation
Semi-anechoic	*A*
Kirishima International Concert Hall (15th row, seat number 7)	*K*
Tsuyama Music Hall (5th row, seat number 9)	*T*

Table 4. Room acoustic measures for conditions *K* and *T*. *A*: A-weighted. *ITDG*: Initial time delay gap between the direct sound and the first reflection. *EDT*: Early decay time. RT_{30}: Reverberation time. *IACC*: Interaural cross correlation. τ_{IACC}: Interaural time difference. W_{IACC}: Width of the interaural cross correlation. *A*-value: Total amplitude of the reflections. *L*: Left channel. *R*: Right channel.

		Room			
		K		*T*	
		All (Flat)	All (*A*)	All (Flat)	All (*A*)
ITDG [ms]	*L*	31		42	
	R	43		62	
EDT [s]		1.8	1.9	2.0	2.0
RT_{30} [s]		2.0	2.0	2.1	2.0
IACC		0.12	0.09	0.11	0.10
τ_{IACC} [ms]		0.083	0.063	0.000	0.000
W_{IACC} [ms]		0.058	0.055	0.052	0.071
A-value	*L*	6.1	5.5	3.8	3.4
	R	5.6	5.2	3.5	3.4

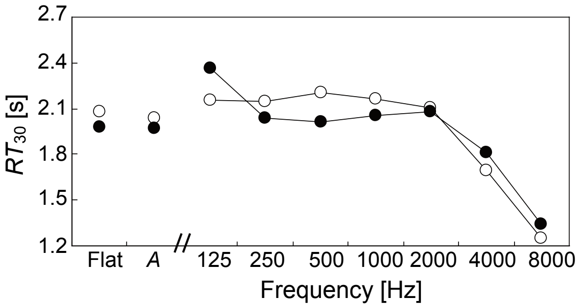

Figure 1. Reverberation time (RT_{30}) in 1/1 octave bands for conditions *K* (●) and *T* (○).

Participants

Fifteen instrumentalists—seven string instrumentalists, four wood wind instrumentalists, and four brass instrumentalists—who perform as members of the same amateur orchestra group of Osaka University participated in the listening experiment. The instrumentalists have each been playing their respective instruments for at least six years.

Procedure

The scale values of timbral brightness for the respective semi-anechoic and reverberant stimuli at equal notes were obtained through Scheffe's paired-comparison test. There were three sessions, and each session consisted of 72 pairs (N (N - 1), N = 9) of stimuli with the same note. The silent interval between the stimuli was 1 s. Each pair was separated by an interval of 4 s. The participants were asked to rate the brightness differences within pairs on a seven point scale, as shown in Figure 2, and asked to fill-in the form during the silent interval between the presentations of each pair.

The nine stimuli with the same note were presented in random order in advance of each session to present the variation range of the brightness within the nine stimuli to the listeners. The stimuli were presented in random order for each participant and for each session. Each session took about 20 min.

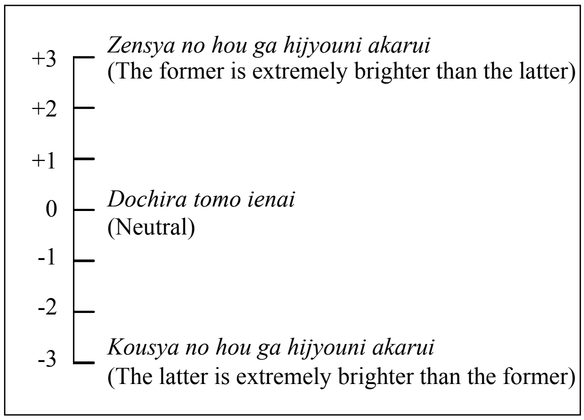

Figure 2. Scale used for listening experiment

The stimuli were sampled at a rate of 44100 Hz with 16-bit resolution and presented dichotically over Sennheiser HD650 headphones to the participants at approximately 60 dB SPL for equal loudness in a soundproof room. A VAIO VGN-FJ11/W notebook computer and M-Audio ProFire Lightbridge 34-in/36-out digital-to-analogue converter were used to reproduce the stimuli. Before the listening experiment, one of the experimenters adjusted the sound reproduction level to 60 dB SPL using a semi-anechoic sound stimulus (A4, *MA*) based on the method of adjustment.

RESULTS

Measured data for one of the 15 participants were eliminated from the dataset before the analysis because the measured scale value of brightness for each stimulus with the larger dynamic level *F* for this participant was lower than each stimulus with the smaller dynamic level *P*. This result contradicts the results for the other 14 participants and that of a previous study showing that the scale value of brightness (*brightness*) normally increases with increasing production level of musical sound (Meyer, 2009).

Table 5 shows the analysis of variance (ANOVA) results obtained by applying the analysis method to Scheffe's paired-comparison test. Each effect excluding the order effect for the experiment results with the lowest note A3 was statistically significant, while the contribution ratio (η^2) for each (≤2.2%) of the combinatorial effect, order effect, and interaction between the order effect and participant was relatively smaller than that for the main effect (≥54%) and/or that for the interaction of the main effect and participant (≥14.9%).

Figure 3 shows the *brightness* for respective semi-anechoic stimuli and their reverberant stimuli at equal note. Each error bar indicates 95% confidence interval. The mean *brightness* ranged between -1.18 and +1.18 for note A3, the mean *brightness* ranged between -1.35 and +1.13 for note A4, while the mean *brightness* ranged between -1.22 and +1.07 for note A5. As a general tendency, the *brightness* increased with increasing dynamic level.

Whether the measured *brightness* differs with *room* was tested for each note and dynamic level. Figure 3 shows that the maximum inter-*room* difference of *brightness* was significant for each note and dynamic level. In extreme cases (between *FT* and *FA* at note A3), the difference in *brightness* was 0.91. These results support the hypothesis that *room* affects the timbral brightness of the musical sound.

Whether the effect of *room* on *brightness* differs with the produced dynamic level was tested for stimuli produced with the same note (A4) but different dynamic levels (*F* and *P*). Figure 3 shows that *brightness* for the stimulus produced with lower dynamic level *P* in *room T* was 0.17 smaller than that for the stimulus in *room K*, while *brightness* for the stimulus produced with the larger dynamic level *F* in *room T* was 0.30 larger than that for the stimulus under *room K*. These results support the hypothesis that the effect of *room* on the timbral brightness of musical sound depends on the produced dynamic level.

Whether the effect of *room* on the *brightness* differs with the produced note was tested for stimuli produced with the same dynamic level (*F*) but different notes (A3 and A5). Figure 3 shows that *brightness* for the stimulus produced at A3 under *room T* was 0.76 larger than that for the stimulus in *room K*, while *brightness* for the stimulus produced at A5 in *room T* was 0.13 smaller than that for the stimulus in *room K*. These results support the hypothesis that the effect of *room* on the timbral brightness of musical sound depends on the produced note.

Table 5. ANOVA results (**: 1% significant level, **: 5% significant level)

(a) Note = A3 (≈220 Hz)

Source	Sum of square	Degree of freedom	η^2 (%)
Main effect	1264.5	8	54.0**
Main effect × Participant	349.8	104	14.9**
Combinatorial effect	46.6	28	2.0**
Order effect	0.1	1	0.0
Order effect × Participant	30.3	13	1.3**
Residual	650.7	854	27.8
Total	2342.0	1008	

(b) Note = A4 (≈440 Hz)

Factor	Sum of square	Degree of freedom	η^2 (%)
Main effect	1826.8	8	65.7**
Main effect × Participant	435.1	104	15.7**
Combinatorial effect	23.9	28	0.9*
Order effect	11.4	1	0.4**
Order effect × Participant	16.2	13	0.6**
Residual	465.7	854	16.8
Total	2779.0	1008	

(c) Note = A5 (≈880 Hz)

Factor	Sum of square	Degree of freedom	η^2 (%)
Main effect	1424.9	8	55.0**
Main effect × Participant	424.7	104	16.4**
Combinatorial effect	44.6	28	1.7**
Order effect	56.2	1	2.2**
Order effect × Participant	30.8	13	1.2**
Residual	610.9	854	23.6
Total	2592.0	1008	

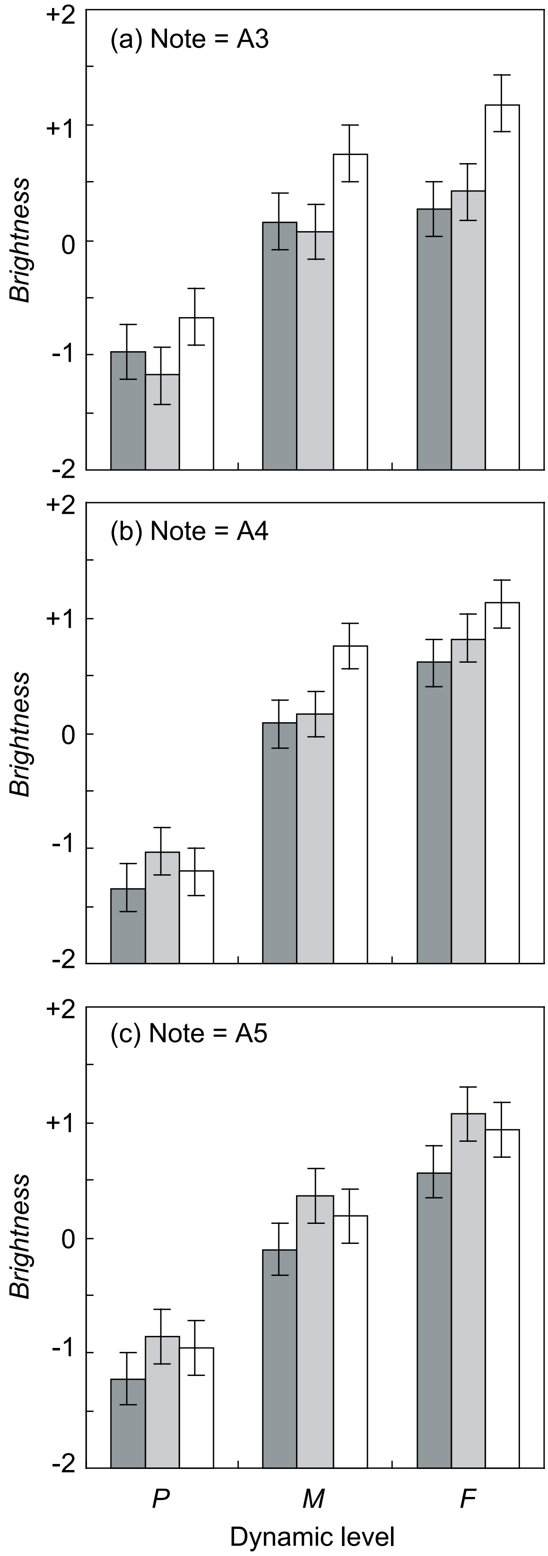

Figure 3. Comparison of effect of room acoustic condition on scale value of timbral brightness of clarinet tones produced at different dynamic levels and different notes. The error bar indicates the 95% confidence interval. ■: room *A*, ■: room *K*, and □: room *T*.

DISCUSSION

The listening experiment demonstrated that the *room* significantly affects the brightness of clarinet tones, while the effect depends on the produced dynamic level and/or note (Table 5 and Figure 3). These results indicate that perceptual brightness in rooms may be described dependent of the acoustic characteristics of the produced musical sound signals. However, it is still unclear which cues exactly contribute to the brightness perceived in rooms. This calls for further detailed investigation.

The clarinet has been indicated to have the largest variation in *brightness* among musical instruments (Meyer, 2009). This study employed the clarinet sound as sound stimuli. The results for different instruments may be not equal to the present results.

The *brightness* normally increases with increasing dynamic level (Meyer, 2009). In this experiment, we produced the stimuli with equal loudness in consideration of the definition of timbre. If the stimuli were presented at different levels of loudness, the results may differ from the present experiment.

In this experiment, the duration of each stimulus was unchanged in order to maintain the perceptual naturalness of the stimuli as much as possible. An experiment with sound stimuli clipped to the same duration may show results unequal to the present results.

The perception of sound stimuli reproduced by a loudspeaker may be different from that reproduced by headphones. To generalize the results, we need to compare the present results with those from further experiments employing loudspeakers and/or further *in situ* experiments.

In this study, as an initial step, just the perceptual brightness was examined because "bright" was assumed to not only be a timbre adjective commonly used by the participants but also a perceptual dimension which is potentially described in relation to the single descriptor of a spectral centroid. Investigations of other scales such as "loudness," "pitch," "reverberance," and "clarity" remain undone and will be addressed in future work.

CONCLUSIONS

The results of this study show that the *room* significantly affects the brightness of clarinet tones, and the effect differs with the produced dynamic level and the produced note. These findings call for a further examination of the interaction between the performing style and the room acoustics with respect to on the brightness of clarinet tones; this will be presented in Part II of this study based on an acoustic analysis.

Studies of the timbre perceived in rooms in relation to both the room acoustic conditions and acoustic characteristics of the produced music sound signals may contribute to a better understanding of the requirements of both the audience and musicians in terms of concert hall acoustics.

ACKNOWLEDGEMENTS

We thank all the instrumentalists who participated in this study for their active cooperation. The authors would like to thank Kazunobu Sagara, Hisashi Kotani, and Yoshihisa Momoi for their helpful suggestions.

REFERENCES

Bech, S., "Timbral aspects of reproduced sound in small rooms. I" *J. Acoust. Soc. Am.* **97**, 1717–1726 (1995)

Bech, S., "Timbral aspects of reproduced sound in small rooms. II" *J. Acoust. Soc. Am.* **99**, 3539–3549 (1996)

Disley, A., and Howard, D., "Spectral correlates of timbral semantics relating to the pipe organ" *TMH-QPSR, KTH* **46**, 25–40 (2004)

Fujii, K., Hotehama, T., Kato, K., Shimokura, R., Okamoto, Y., Suzumura, Y., and Ando, Y., "Spatial distribution of acoustical parameters in concert halls: comparison of different scattered reflections" *J. Temporal Des. Arch. Environ.* **4**, 59–68 (2004)

Goto, M., Hashiguchi, H., Nishimura, T., and Oka, R., "RWC Music Database: Music Genre Database and Musical Instrument Sound Database" *Proc. 4th International Conference on Music Information Retrieval* (ISMIR 2003), 229–230 (2003)

Hikichi, T., and Miyoshi, M., "Sound timbre control using estimates of room resonance modes" *Acoust. Sci. Tech.* **27**, 257–263 (2006)

Hotehama, T., Sato, S., and Ando, Y., "Dissimilarity judgments in relation to temporal and spatial factors for the sound field in an existing hall" *J. Sound and Vib.* **258**, 429–441 (2002)

Kato, K., Nagao, T., Yamanaka, T., Kawai, K., and Sakakibara, K., "Study on effect of room acoustics on timbral brightness of clarinet tones. Part II: an acoustic interpretation and synthesis of analytical results" *Proceedings of the 20th International Congress on Acoustics, Sydney*, (2010)

Kawai, K., Kojima, T., Hirate, K., and Yasuoka, M., "Personal evaluation structure of environmental sounds: experiments of subjective evaluation using subjects' own terms" *J. Sound and Vib.* **277**, 523–533 (2004)

Kuttruff, H., *Room Acoustics (fourth edition)* (Spon Press, London, 2000) Chapter 7

Marozeau, J., Cheveigé, A., McAdams, S., and Winsberg, W., "The dependency of timbre on fundamental frequency" *J. Acoust. Soc. Am.* **114**, 2946–2957 (2003)

Marozeau, J., and Cheveigé, A., "The effect of fundamental frequency on the brightness dimension of timbre" *J. Acoust. Soc. Am.* **121**, 383–387 (2007)

Meyer, J., "Vibrato Sounds in Large Halls" *Proc. Stockholm Music Acoust. Conf.*, 117–121 (1993)

Meyer, J., *Acoustics and the Performance of Music (fifth edition)* (Springer, 2009) pp. 347–387

Nanba, S., and Kuwano, S., *Method of Psychological Measurement for Hearing Research* (Corona Publishing Co., Ltd., Tokyo, 1998) pp. 126 (in Japanese)

Schubert, E., and Wolfe, J., "Does timbral brightness scale with frequency and spectral centroid?" *Acta Acustica united with Acustica* **92**, 820–825 (2006)

Ueno, K., Kato, K., and Kawai, K., "Effect of room acoustics on musicians' performance. Part I: experimental investigation with a conceptual model" *Acta Acustica united with Acustica* **96**, 505–515 (2010)

Zahorik, P., "Perceptually relevant parameters for virtual listening simulation of small room acoustics" *J. Acoust. Soc. Am.* **126**, 776–791 (2009)