Effect of room acoustics on timbral brightness of clarinet tones: Experimental investigation with two binaural room impulse responses

Kosuke Kato\textsuperscript{1,}\textsuperscript{*}, Tsubasa Nagao\textsuperscript{2}, Toshio Yamanaka\textsuperscript{3}, Keiji Kawai\textsuperscript{4} and Ken-Ichi Sakakibara\textsuperscript{5}

\textsuperscript{1}Office for University-Industry Collaboration, Osaka University, Yamadaoka 2–8, Suita, 565–0871 Japan
\textsuperscript{2}School of Engineering, Osaka University, Yamadaoka 2–1, Suita, 565–0871 Japan
\textsuperscript{3}Graduate School of Engineering, Osaka University, Yamadaoka 2–1, Suita, 565–0871 Japan
\textsuperscript{4}Graduate School of Science and Technology, Kumamoto University, Kurokami 2–39–1, Kumamoto, 860–8555 Japan
\textsuperscript{5}School of Psychological Science, Health Sciences University of Hokkaido, Ainosato 2–5, Kita-ku, Sapporo, 002–8072 Japan

(Received 20 November 2013, Accepted for publication 3 July 2014)

Abstract: This study is aimed at investigating how room acoustics affect the timbral brightness of clarinet tones. For semi-anechoic stimuli, nine natural clarinet tones produced at three different dynamic levels and three different notes were used. Reverberant stimuli were generated by convolving each semi-anechoic tone with two different binaural room impulse responses. The stimuli were presented dichotically over headphones to the fifteen participants at equal loudness. The scale values of timbral brightness for each stimulus at equal notes were obtained through Scheffe’s paired-comparison test. The results showed that timbral brightness was significantly varied depending on room acoustic conditions. The results also showed that the spectral centroid of tones produced at equal dynamic levels and equal notes was varied depending on room acoustic conditions. The variation range was equivalent to 48–52% of the overall range of that caused by varying both room acoustic conditions and dynamic levels at each note. Timbral brightness was found to linearly increase with the increase in the spectral centroid, irrespective of room acoustics. The slope of timbral brightness corresponding to the spectral centroid was about 1.5–1.6 times steeper for the reverberant tones than for the semi-anechoic tones.

Keywords: Timbre, Tone quality, Brightness, Room acoustics, Reverberation, Spectral centroid, Clarinet

PACS number: 43.55.Hy, 43.66.Jh, 43.75.Cd, 43.75.Pq [doi:10.1250/ast.35.300]

1. INTRODUCTION

The aim of this study is to investigate how room acoustics affect the timbral brightness of clarinet tones produced at different dynamic levels and different notes. The understanding of such a room acoustic effect can be useful to sound engineers and musicians, for example, to convey desired musical timbre to the audience in the room.

In order to understand such a room acoustic effect, it will be useful to examine the relationship between the timbral brightness and acoustic descriptor(s) of clarinet tones. It is well known that timbral brightness increases with increasing power at high frequencies. The distribution in the power spectrum can be simply quantified using the spectral centroid.

In previous studies, several researchers showed that timbral brightness positively correlates with either the spectral centroid [1] or the spectral centroid divided by fundamental frequency [2]. Also, several studies revealed that one of the dimensions of timbre space positively correlates with spectral centroid [3–6] or the spectral centroid corrected as a function of fundamental frequency [7].
As for the investigation of the influence of room acoustics on the timbral brightness of musical sound, only one attempt has been made [2]. It was shown that a more reverberant pipe organ sound was perceived as less bright, associated with the decrease in the spectral centroid divided by fundamental frequency [2].

In order to further understand the influence of room acoustics on the timbral brightness of musical sounds, we examine the answer to each of the following three research questions experimentally.

Q1 How do room acoustic conditions affect the timbral brightness of clarinet tones produced at different dynamic levels and different notes?
Q2 How do room acoustic conditions affect the spectral centroid of clarinet tones produced at different dynamic levels and different notes?
Q3 How do room acoustic conditions affect the relationship between the timbral brightness and spectral centroid of clarinet tones produced at different dynamic levels and different notes? Does timbral brightness positively correlate with the spectral centroid of clarinet tones, irrespective of room acoustics?

2. METHOD

2.1. Confirmation of Timbral Brightness as One of the Timbre Adjectives Commonly Used by Potential Participants of Listening Experiment

In order to subjectively evaluate the timbre of musical sounds using adjectives, a timbre adjective 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 [8]. The former better maintains the generality of the experimental results, while the latter better avoids 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 by potential participants in the following listening experiment were examined.

Forty Japanese-speaking instrumentalists—sixteen string instrumentalists, thirteen woodwind instrumentalists, and eleven brass instrumentalists—performing as members of the same amateur orchestral group at Osaka University took part in a questionnaire survey.

A list of eighty-four adjectives, including forty pairs of adjectives used in previous studies [9] and four adjectives chosen on the basis of one of the experimenter’s experience, was presented to each participant. Each participant was asked to judge his/her frequency of use of each adjective into one of 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 of each adjective greatly depends on the instrumental group of participants, 90% of the participants in each instrumental group evaluated ten 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).

Thus, “bright” was confirmed to be one of the timbre adjectives commonly used by potential participants in the listening experiment.

2.2. Listening Experiment to Measure Timbral Brightness

Semi-anechoic stimuli

For semi-anechoic stimuli, nine natural clarinet tones of three different notes (A3 ∼ 220 Hz, A4 ∼ 440 Hz, and A5 ∼ 880 Hz) produced at three different dynamic levels (P, M, and F), as shown in Table 1, were extracted from the RWC music database, which stores musical instrument sounds [10]. The duration of each stimulus was between 2.1 and 3.1 s, as shown in Table 2.

Clarinets 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. It was reported that room acoustics affect the perception of the vibrato sound [11]. 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 has one of the largest brightness ranges among instrumental sounds. It was reported that the produced dynamic level affects the timbral brightness of a musical tone [12]. It was also indicated that the clarinet can produce the softest pianissimo (down to 57 dB) among all wind instruments and a high power fortissimo (up to 106 dB) [12]. This results in a dynamic range of a breadth that is rarely found in any other instrument. 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.

![Table 1 Dynamic levels produced and abbreviations.](https://example.com/table1)

<table>
<thead>
<tr>
<th>Dynamic level</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piano</td>
<td>P</td>
</tr>
<tr>
<td>Mezzo</td>
<td>M</td>
</tr>
<tr>
<td>Forte</td>
<td>F</td>
</tr>
</tbody>
</table>
Table 2  Duration of each semi-anechoic stimulus.

<table>
<thead>
<tr>
<th>Note</th>
<th>Dynamic level</th>
<th>Duration [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3</td>
<td>P</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>2.4</td>
</tr>
<tr>
<td>A4</td>
<td>P</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>2.1</td>
</tr>
<tr>
<td>A5</td>
<td>P</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Table 3  Abbreviation of room acoustic conditions.

<table>
<thead>
<tr>
<th>Room acoustic condition</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-anechoic</td>
<td>A</td>
</tr>
<tr>
<td>Kirishima International Concert Hall</td>
<td>K</td>
</tr>
<tr>
<td>(15th row, seat number 7)</td>
<td></td>
</tr>
<tr>
<td>Tsuyama Music Hall</td>
<td>T</td>
</tr>
<tr>
<td>(5th row, seat number 9)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4  Room acoustic measures calculated according to ISO 3382-1 for room acoustic conditions K and T.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Room K</th>
<th>Room T</th>
<th>Difference</th>
<th>JND[^1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDT [s]</td>
<td>1.62</td>
<td>2.07</td>
<td>24.7%</td>
<td>5%</td>
</tr>
<tr>
<td>$T_{30}$ [s]</td>
<td>2.05</td>
<td>2.05</td>
<td>0.0%</td>
<td>5%</td>
</tr>
<tr>
<td>$T_{5}$ [s]</td>
<td>0.129</td>
<td>0.146</td>
<td>0.017</td>
<td>0.010</td>
</tr>
<tr>
<td>IACC</td>
<td>0.360</td>
<td>0.330</td>
<td>0.030</td>
<td>0.075</td>
</tr>
</tbody>
</table>

Reverberant stimuli

For reverberant stimuli, eighteen tones were generated by convolving each semi-anechoic tone with two different binaural room impulse responses (BRIRs), which were collected at a different seat position in two different medium-sized concert halls, as shown in Table 3.

Table 4 lists some of the room acoustic measures ($EDT$, $T_{30}$, $T_{5}$, and IACC) for describing the room acoustic conditions of each room, as calculated in accordance with ISO 3382-1 [13]. In order to compress the information, the average across each measure from each of six-octave frequency ranges, 125, 250, 500, 1,000, 2,000, and 4,000 Hz bands, is shown. The difference between conditions $K$ and $T$ in each measure, and the just noticeable difference (JND) ([14] and references therein) for each measure are also shown for reference.

The two BRIRs were chosen on the basis of the results of a preliminary listening experiment carried out by one of the experimenters as follows.

First, twenty-two BRIRs collected at twenty-two representative seat positions in the above mentioned two concert halls were convolved with each of the nine semi-anechoic stimuli. One of the experimenters possessed those BRIRs. The spatial distribution for acoustical parameters in these two concert halls was previously analyzed in detail by a group including one of the authors [15].

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 by a signal processing technique.

Third, the experimenter wore HD650 headphones connected to the headphone jack of a VAIO VGN-FJ11/W notebook computer. The sound reproduction level was then adjusted to a comfortable level by the experimenter. The experimenter then randomly reproduced and compared each pair of the twenty-two reverberant stimuli generated by convolving the twenty-two BRIRs with a semi-anechoic clarinet tone of note A4 produced at dynamic level $M$.

Finally, the experimenter chose two BRIRs with which the clarinet tone was perceived by the experimenter as the brightest and the least bright among the twenty-two BRIRs, respectively.

The choice of only these two BRIRs limits the duration of the listening experiment per participant to within approximately one hour, which decreases the load on the participants. Two BRIRs are considered to be sufficient to investigate how much variation in the timbral brightness of clarinet tones is caused by convolving these two BRIRs in comparison with that caused by varying the dynamic level or fundamental frequency.

Participants

Fifteen instrumentalists—seven string instrumentalists, four woodwind instrumentalists, and four brass instrumentalists—who perform as members of the same amateur orchestral 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 were obtained through Scheffe’s paired-comparison test. Hereafter, the scale value of timbral brightness is denoted as brightness($i$, $j$, $k$), where

\[
i = \text{room acoustic condition (room)} (A, K, \text{ and } T),\]

\[
j = \text{produced dynamic level} (P, M, \text{ and } F),\]

\[
k = \text{produced note} (A3, A4, \text{ and } A5).
\]

There were three sessions, and each session consisted of seventy-two 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 Fig. 1, and asked to fill-in the form during the silent interval between the presentations of each pair.

The stimuli were sampled at a rate of 44,100 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 sound-proof room. A VAIO VGN-FJ11/W notebook computer and an 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 (A, M, A4) by the method of adjustment.

2.3. Analysis of Spectral Centroid

The spectral centroid \( f_c(i,j,k) \) of clarinet tones for room \( i \), dynamic level \( j \), and note \( k \) was analyzed acoustically for amplitude spectrum \( A_m(f_m) \). In this study, \( f_c \) was defined as \( \sum f_m A_m / \sum A_m \), where \( m \) was the discrete Fourier transform coefficient, \( f_m \) was its frequency on a linear scale, and \( A_m \) was its amplitude on a linear scale. In this study, \( f_c(i,j,k) \) for the binaural stimuli was defined as the average spectral centroid of signals of the left and right channels formulated as

\[
f_c(i,j,k) = \frac{f_{c, left}(i,j,k) + f_{c, right}(i,j,k)}{2}.
\]  

2.4. Relationship between Timbral Brightness and Measured Spectral Centroid

The relationship between brightness \((i,j,k)\) and measured \( f_c(i,j,k) \) was investigated by testing the following hypotheses in three steps, H1 to H3.

[H1] The brightness \((i,j,k)\) can be described by a linear equation, setting \( f_c \) as a factor, as

\[
brightness(i,j,k) = a_0(i) + a_1(i) \cdot f_c(i,j,k),
\]

where \( a_0(i) \) is the constant term for room \( i \), and \( a_1(i) \) is the linear coefficient for room \( i \). Furthermore, \( [H2] \) either \( a_0(i) \) or \( a_1(i) \) in Eq. (5) varies with the room; \( [H3] \) each of \( a_0(i) \) and \( a_1(i) \) in Eq. (5) correlates with some of the acoustic measures of the room.

3. RESULTS

3.1. Timbral Brightness

Before the analysis, measured data for one of the fifteen participants were eliminated from the dataset because the measured scale value of brightness for each stimulus with the higher dynamic level \( F \) for this participant was lower than each stimulus with the lower dynamic level \( P \). This result contradicts the results for the other fourteen participants and that of a previous study showing that the scale value of brightness (brightness) normally increases with increasing production level of musical sound [12].

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 experimental results with the lowest note A3 was statistically significant, while the contribution ratio (\( \eta^2 \)) for each (\( \leq 2.2\% \)) of the combinatorial effect,
indicates that the higher the notes, the smaller the range. Each of these ranges was equivalent to 39%, 27%, and 22% of the total range of \(\text{brightness}(i, j, k)\) among all nine stimuli at each note. Whether the \textit{room} made \(\text{brightness}(i, j, k)\) lower or higher was dependant on both note and dynamic level.

The range of \(\text{brightness}(i, j, k)\) depending on dynamic level produced at an equal note for each \textit{room} was the greatest at 2.33 \((p < 0.01)\) for note A4, followed by 1.92 \((p < 0.01)\) for note A5 and 1.85 \((p < 0.01)\) for A3. Each of these ranges was equivalent to 94%, 84%, and 78% of the total range of \(\text{brightness}(i, j, k)\) among all the nine stimuli at each note. As a general tendency, \(\text{brightness}(i, j, k)\) was the greatest for dynamic level \(F\), followed by dynamic level \(M\) and dynamic level \(P\). This indicates that the higher the dynamic level, the greater the \(\text{brightness}(i, j, k)\).

### 3.2. Measured Spectral Centroid

Figure 3 shows measured \(f_c(i, j, k)\) for nine stimuli at each of three different notes, as formulated in Eq. (4). The range of \(f_c(i, j, k)\) among all nine stimuli at each note was the largest at 1.60 kHz for note A3, followed by 1.56 kHz for note A4 and 1.24 kHz for note A5. This indicates that the higher the notes, the smaller the range of \(f_c(i, j, k)\).

The range of \(f_c(i, j, k)\) depending on \textit{room} produced at the same dynamic level and equal note was the largest at 831 Hz for note A3, followed by 749 Hz for note A4 and 619 Hz for A5. This again indicates that the higher the notes, the smaller the range of \(f_c(i, j, k)\) among all nine stimuli at each note. Whether the \textit{room} made \(f_c(i, j, k)\) lower or higher was dependant on both note and dynamic level.

The range of \(f_c(i, j, k)\) depending on dynamic level produced at an equal note for each \textit{room} was the greatest at 1.20 kHz for note room A, followed by 991 Hz for room \(T\) and 794 Hz for room \(K\). Each of these ranges was equivalent to 77%, 64%, and 51% of the overall range of \(f_c(i, j, k)\) among all nine stimuli at each note. As a general tendency, \(f_c(i, j, k)\) was the highest for dynamic level \(F\), followed by dynamic level \(M\) and dynamic level \(P\). This indicates that the higher dynamic level, the greater the \(f_c(i, j, k)\).

### 3.3. Relationship between Timbral Brightness and Measured Spectral Centroid

Table 6 lists the results of a linear regression, setting \(f_c(i, j, k)\) as a factor, subjected to all twenty-seven conditions of \(\text{brightness}(i, j, k)\) (see Eq. (5)). In this model, a linear relationship was observed, with a coefficient of determination \((R^2)\) of 0.563 and a corrected Akaike’s Information Criterion (AICc) of 50.3.
Table 7 Results of effect test for the adaptation of a model with \( f_c(i, j, k) \) and room set as factors, subjected to all twenty-seven conditions of brightness\((i, j, k)\). \( R^2 = 0.897 \) and AICc = 17.0. DF: degree of freedom. MS: mean square.

<table>
<thead>
<tr>
<th>Factor</th>
<th>DF</th>
<th>MS</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_c )</td>
<td>1</td>
<td>15.2</td>
<td>189</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>room</td>
<td>2</td>
<td>6.00</td>
<td>37.4</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 8 Results of effect test for the adaptation of a model with \( f_c(i, j, k) \), room, and their cross-effect \((f_c(i, j, k) \times \text{room})\) set as factors, subjected to all twenty-seven conditions of brightness\((i, j, k)\). \( R^2 = 0.934 \) and AICc = 11.9.

<table>
<thead>
<tr>
<th>Factor</th>
<th>DF</th>
<th>MS</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_c(i, j, k) )</td>
<td>1</td>
<td>15.6</td>
<td>278</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>room</td>
<td>2</td>
<td>5.00</td>
<td>44.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( f_c(i, j, k) \times \text{room} )</td>
<td>2</td>
<td>0.663</td>
<td>5.91</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 9 Single variable linear regression for each room. Variables \( a_0(i) \) and \( a_1(i) \) are regression constants: the constant term and linear term, respectively. \( R^2 \): coefficient of determination. AICc: corrected Akaike’s Information Criterion.

<table>
<thead>
<tr>
<th>room</th>
<th>( a_0(i) )</th>
<th>( a_1(i) )</th>
<th>( R^2 )</th>
<th>AICc</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, K, and T</td>
<td>(-2.84)</td>
<td>(1.46 \cdot 10^{-3})</td>
<td>0.943</td>
<td>4.61</td>
</tr>
<tr>
<td>K</td>
<td>(-2.79)</td>
<td>(2.20 \cdot 10^{-3})</td>
<td>0.933</td>
<td>7.39</td>
</tr>
<tr>
<td>T</td>
<td>(-3.14)</td>
<td>(2.29 \cdot 10^{-3})</td>
<td>0.921</td>
<td>11.2</td>
</tr>
</tbody>
</table>

Table 7 lists the results for a model, setting \( f_c(i, j, k) \) and room as factors, subjected to the twenty-seven conditions of brightness\((i, j, k)\). In this model, a better fitting was observed, with \( R^2 \) of 0.897 and AICc of 17.0. Here, the resulting \( R^2 \) was higher and the obtained AICc was lower than the value listed in Table 6, which indicates that the model in Table 7 is better than the model in Table 6.

Table 8 lists the results for a model, setting \( f_c(i, j, k) \), room, and their cross-effect \((f_c(i, j, k) \times \text{room})\) as factors, subjected to the twenty-seven conditions of brightness\((i, j, k)\). In this model, a better fitting was observed, with \( R^2 \) of 0.934 and with AICc of 11.9. Here, the resulting \( R^2 \) was higher and the obtained AICc was lower than the values listed in Tables 6 and 7, which indicates that the model in Table 8 is better than the models in Tables 6 and 7. This shows that not only \( f_c(i, j, k) \) but also room and their cross-effect \((f_c(i, j, k) \times \text{room})\) significantly contribute to brightness\((i, j, k)\).

Table 9 lists the results of a linear regression, setting \( f_c(i, j, k) \) as a factor, subjected to nine conditions of brightness\((i, j, k)\) in each room. These models are also illustrated in Fig. 4. In each model, a significant linear relationship was observed, with a coefficient of determination \( (R^2) \) ranging between 0.921 and 0.943 and with an AICc between 4.61 and 11.2. Here, each of the resulting \( R^2 \) was higher and each of the obtained AICc was lower than the values listed in Tables 6, 7, and 8, which indicates that each model in Table 9 is better than the models in Tables 6, 7, and 8. These results show that brightness in each room linearly increases with \( f_c(i, j, k) \).
4. DISCUSSION

As noted in Introduction, the goal of the present study was to investigate how room acoustics affect the timbral brightness of clarinet tones produced at different dynamic levels and different notes. In each following subsection, we discuss the answer to each of the three research questions posed in Introduction, in comparison with the findings in published studies.

4.1. Effect of Room Acoustic Conditions on Timbral Brightness of Clarinet Tones Produced at Different Dynamic Levels and Different Notes

Given the results in Fig. 2 and those of [2], it can be concluded that the timbral brightness of musical tones can be significantly varied depending on room acoustic conditions. Figure 2 provides an example of quantified ranges of timbral brightness of musical tones depending on room acoustic conditions relative to that depending on dynamic level.

Why does Fig. 2 show a smaller range of timbral brightness depending on room acoustic conditions for the higher notes? Why does Fig. 2 show that the room acoustics conditions made clarinet tones either brighter or less bright depending on note or dynamic level. These two findings may be observed because room acoustic conditions affected the relationship between the timbral brightness and spectral centroid of clarinet tones, as will be described in Subsect. 4.3.

The results in Fig. 2 confirmed that the timbral brightness of clarinet tones increased with a higher dynamic level. This result is in accordance with the indication that the produced dynamic level affects the timbral brightness of musical tone [12].

4.2. Effect of Room Acoustic Conditions on Spectral Centroid of Clarinet Tones Produced at Different Dynamic Levels and Different Notes

Taking into account the results in Fig. 3 and those of [2], it can be said that the spectral centroid of musical tones can be varied depending on room acoustic conditions. The results in Fig. 3 are quantified ranges of spectral centroid of musical tones depending on room acoustic conditions relative to that depending on dynamic level.

Why is a smaller range of spectral centroid observed in Fig. 3 for the reverberant rooms? Why does Fig. 3 show that the spectral centroid of reverberant signals is lower than that of semi-anechoic tones? The latter phenomenon is in accordance with the results obtained in [2], which showed that the value of the spectral centroid divided by fundamental frequency for a reverberant pipe organ sound is lower than those for other less reverberant pipe organ sounds. In order to answer these two questions, the relationship between the distribution in the power spectrum of the musical tones and the frequency-response characteristic of BRIRs, which affects the spectral centroid of the convoluted sound signal, should be examined. This is an issue that invites further work.

The results in Fig. 3 confirmed that the spectral centroid of clarinet tones increased with a higher dynamic level. This result is in agreement with those reported in [12] and references therein that indicate the large difference in spectral slope between the tones produced at a

Table 10 presents the correlation matrix of variables $a_0(i)$, $a_1(i)$, and room acoustic measures listed in Table 4. No significant correlation between the room acoustic measures and either $a_0(i)$ or $a_1(i)$ was found.

### Table 10: Correlation matrix of variables $a_0(i)$, $a_1(i)$, and room acoustic measures listed in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>$a_0(i)$</th>
<th>$a_1(i)$</th>
<th>EDT</th>
<th>$T_{30}$</th>
<th>$T_S$</th>
<th>IACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0(i)$</td>
<td>1</td>
<td>-0.47</td>
<td>-0.18</td>
<td>-0.38</td>
<td>-0.28</td>
<td>0.34</td>
</tr>
<tr>
<td>$a_1(i)$</td>
<td>1</td>
<td>0.95</td>
<td>1.00</td>
<td>0.98</td>
<td>-0.99</td>
<td></td>
</tr>
<tr>
<td>EDT</td>
<td>1</td>
<td>0.98</td>
<td>0.99</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_{30}$</td>
<td>1</td>
<td>1</td>
<td>-1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_S$</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IACC</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
lower dynamic level and the tones produced at a higher
dynamic level.

4.3. Effect of Room Acoustic Conditions on the
Relationship between Timbral Brightness and
Spectral Centroid of Clarinet Tones Produced at
Different Dynamic Levels and Different Notes
Considering the results in Tables 6, 7, 8, and 9 and in
Fig. 4, the timbral brightness of clarinet tones linearly
increases with the increase in the spectral centroid,
irrespective of room acoustics. These results support our
hypothesis H1 and are in agreement with the previous
results that showed that spectral centroid (or spectral
centroid divided by fundamental frequency) positively
correlates with either the timbral brightness of musical
sounds [1,2] or one of the dimensions of timbre space
[3–6]. This finding suggests that we can linearly interpola-
tethe timbral brightness of clarinet sounds simply by
analyzing the spectral centroid of signals recorded there,
as long as the room acoustic condition is constant. This
may be a useful result for researchers, sound engineers, and
musicians who are interested in the timbre of musical
sounds in a given room.

In this study, the slope of timbral brightness corre-
sponding to the spectral centroid was found to be 1.5–1.6
times steeper for the reverberant tones than that for the
semi-anechoic tones. This supports our hypothesis H2. No
similar finding has been reported in the literature, to the
best of our knowledge. This finding suggests that the
timbral brightness under a certain room acoustic condition
can be described not only by the spectral centroid but also
by the other acoustic cues of the room acoustic condition.

The results in Table 10 do not support our hypothesis
H3. This may be mainly because the number of data points
was only three. It is still unclear which cues contribute to
the timbral brightness perceived in rooms because of the
limited number of BRIRs employed in this study. This calls
for further investigation with greater variations of room
acoustic conditions.

4.4. Dependence of the Finding of This Study on
Logarithmic Transformation of Spectral Centroid
When each of $f_{c,\text{left}}$, $f_{c,\text{right}}$, and $f_c$ was transformed to
a logarithmic ($\log_{10}$) scale, the $R^2$ values became higher
($R^2 = 0.958$) in room A but lower in rooms $K$ ($R^2 = 0.878$)
and $T$ ($R^2 = 0.903$) than those for the model described in
Table 9 and Fig. 4. Therefore, the logarithmic transfor-
mation of the spectral centroid likely improves or worsens
the model used to describe timbral brightness.

5. CONCLUSIONS
In order to investigate how room acoustics affect the
timbral brightness of clarinet tones, we experimentally
examined (1) the effect of room acoustic conditions on
timbral brightness, (2) the effect of room acoustic
conditions on the spectral centroid, and (3) the effect of
room acoustic conditions on the relationship between
timbral brightness and the spectral centroid, for the clarinet
tones produced at different dynamic levels and different
notes.

The results of this study lead to the following
conclusions.
1. Depending on room acoustic conditions, the timbral
brightness of clarinet tones produced at equal dynam-
ic levels and equal notes was significantly varied.
The variation range was equivalent to 22–39% of the
overall range of variation caused by varying both
room acoustic conditions and dynamic levels at equal
notes. A smaller variation range was observed for the
higher notes than for the lower notes.
2. The spectral centroid of clarinet tones produced at
equal dynamic levels and equal notes varied depend-
ing on room acoustic conditions. The variation range
was equivalent to 48–52% of the overall range of
variation caused by varying both room acoustic
conditions and dynamic levels at equal notes.
3. Timbral brightness was found to increase linearly
with an increase in the spectral centroid, irrespective
of room acoustics. The slope of timbral brightness
corresponding to the spectral centroid was 1.5–1.6
times steeper for the reverberant tones than that for the
semi-anechoic tones.

These findings are expected to be useful, for example,
for synthesizer manufacturers, sound effector manufactur-
ers, room acoustic engineers, public address (PA) engi-
neers, and musicians as means of realizing practical
applications such as musical sound synthesis, effector
control, room acoustic design, and adjustment of musical
performance depending on room acoustic conditions.

ACKNOWLEDGEMENTS
The authors thank all the instrumentalists who partici-
pated in this study for their willing cooperation. The
authors thank Kazunobu Sagara, Hisashi Kotani, and
Yoshi-hisa Momoi for their helpful suggestions. The
authors also thank the reviewers for their helpful com-
ments and corrections. Portions of this work were
presented in “Study on the effect of room acoustics on
timbral brightness of clarinet tones. Part I: subjective
evaluation through a listening experiment,” Proceedings of
the 20th International Congress on Acoustics, Sydney,
P617 (2010) and “Study on effect of room acoustics on
timbral brightness of clarinet tones. Part II: an acoustic
interpretation and synthesis of analytical results,” Proceed-
ings of the 20th International Congress on Acoustics,
Sydney, P619 (2010).
REFERENCES


