Investigation of the Relation Between $(\tau_e)_{min}$ and Operatic Singing with Different Vibrato Styles

Kosuke Kato[1,a)], Takatsugu Hirawa[2], Keiji Kawai[1], Takashi Yano[1], and Yoichi Ando[3]

[1]Graduate School of Science and Technology Kumamoto University, 2-39-1 Kurokami, Kumamoto 860-8555 Japan

[2]Faculty of Education and Master's Course in Education Kumamoto University, Japan

[3]Professor Emeritus, Kobe University, Japan

This study investigates the relationship between operatic singing and opera house acoustics in terms of the minimum value of the effective duration of the running autocorrelation function $(\tau_e)_{min}$ of voice signals. It has been shown that $(\tau_e)_{min}$ of source signal is closely correlated with the subjective response of both listeners and performers to the temporal conditions of sound fields [Y. Ando: "Architectural Acoustics", AIP/Springer-Verlag, New York (1998)]. In order to examine the variability of $(\tau_e)_{min}$ by means of variation in musical expressivity, we analyzed the steady-state part of anechoic recordings of variably pitched vowels sung with different subjective degrees of vibrato for ten trained singers performing operatic-style sung vowels. The results showed that $(\tau_e)_{min}$ values can be varied by an appropriate selection of the subjective degree of vibrato, tone pitch, and vowel selection. Significant findings are that (1) the contribution of the subjective degree of vibrato to the $(\tau_e)_{min}$ value was typically more significant; $(\tau_e)_{min}$ decreased with increasing subjective degree of vibrato; and (2) the contribution of each musical expression to the value of $(\tau_e)_{min}$ depended greatly on the individual singers.

Keywords: autocorrelation function, operatic singing, opera house, vowel, vibrato, individual singer variation

1.INTRODUCTION

Operatic singers are required to produce both "clear" and "reverberant" voice qualities in opera houses with a wide variety of room acoustical conditions; here, these subjective attributes are considered to be in opposition. Some singers tend to retain their own interpretation styles that have been developed and established during long-term training in lesson rooms. Other singers attempt to adapt their interpretation styles to suit the acoustical conditions of the sound fields of a given performance hall. A cellist termed the latter approach a "lifeline of music" wherein the room acoustic was considered as a second musical instrument, and asserted that "this is difficult but beneficial to be achieved" (Tsutsumi, 2005). In order to discuss the latter approach with respect to the relationship between operatic singing and a given opera house, we define the concept of "Blending of operatic singing with a given opera house." Our aim is to completely achieve the preferred conditions for both listeners and performers. This concept includes both the efforts by operatic singers to adapt their

performance to a given opera house and the efforts by opera house acousticians to adjust the sound fields.

A number of systematic studies on singing voice acoustics, speech acoustics, concert hall acoustics, and opera house acoustics have been independently conducted in each research field thus far; however, there are not many interdisciplinary studies on the relevance of room acoustic parameters to the performers. Meanwhile, in order to assist the onstage musicians by using stage reflections, several subjective experiments have been performed using both simulated sound fields and existing halls (Marshall et al., 1978; Gade, 1989a; Gade, 1989b; Chiang et al., 2003; Ueno and Tachibana, 2003; Ueno and Tachibana, 2005; Ueno et al., 2005). Moreover, knowledge on interdisciplinary works encompassing sound source signal acoustics and room acoustics is scarce. With regard to the spatial characteristics of the source signals of the musical sounds, Marshall and Meyer (1985) measured the directivity and auditory impression of singers through experiments with simulated reflections. Otondo and Rindel (2004) investigated the influence of the directivity of musical instruments on the distribution of the acoustical parameters of a room. The preferred self-to-other ratio (SOR) for choir singers was investi-

a) Now at Center for Advanced Science and Innovation, Osaka University, Suita, Osaka 565-0871, Japan

gated to determine the adequate designs of the stages and of choral formations (Ternström, 1999; Ternström et al., 2005). In the frequency-domain analysis of the source signals of the musical sounds, Ternström (1989) analyzed the long-time-average-spectrum (LTAS) of different choirs in different rooms; they reported that room acoustics significantly influenced both the LTAS shape and the sound level of singing voice signals. To improve speech intelligibility in reverberant environments by the preprocessing of a speech source signal, Arai et al. (2002) proposed a technique for suppressing the steady-state portions of speech, and its effectiveness in improving speech intelligibility was confirmed through listening experiments (Arai et al., 2002; Hodoshima et al., 2004; and Hodoshima et al., 2005).

To describe the relationship between the temporal characteristics of sound source signals and those of room acoustics, the decay characteristic of the autocorrelation function (ACF) of a sound source signals has been investigated. It has been shown to be a fundamental parameter in the analysis of the subjective preference of sound fields, and it represents the temporal coherence (Ando, 1998; Ando et al., 2000; Ando, 2002; Sakai et al., 2004). The minimum value of the effective duration of the running ACF (r-ACF) $(\tau_e)_{min}$ is a temporal parameter of the given sound source that is closely related to the preferred temporal conditions for listeners (Ando, 1989; Ando, 1998; Ando et al. 2000; Ando 2002; and Kuroki et al., 2004) and performers (Nakayama, 1984; Sato et al., 2000; Noson et al., 2000; and Noson et al., 2002). In other words, the preferred condition of the temporal parameters of the sound field can be determined from $(\tau_e)_{min}$ values of a specific source signal and conversely, the preferred $(\tau_e)_{min}$ value of the source signal can also be estimated by the temporal parameters of a given sound field. This role of τ_e in room acoustics is briefly introduced in APPENDIX A. The effective duration τ_e also represents the total amount of the randomness of the sound signals due to fluctuating factors such as vibrato, intonation, and jitter (Sundberg, 1987; Titze, 1994; Prame, 1994; Prame, 1997). The effective duration τ_e decreases with increasing randomness of the fluctuations.

If musicians and composers are not familiar with the meaning of the room acoustical parameter for sound signals, it would be important to describe the τ_e value in relation to the musical score and musical expressivity. Previous studies measured $(\tau_e)_{min}$ for several types of music signals, although the main purpose of these studies was to clarify the relationship between $(\tau_e)_{min}$ of a sound source signal and the psychological response to the sound field of a performance hall (Ando et al.,

1989; Ando, 1998; and Kuroki et al., 2004.) In an investigation of the variation in the r-ACF for musical signals relating to performing style-including tempo, articulation, and damper pedaling-these were shown to be the main expressive factors for controlling τ_e of the music from a piano (Taguti and Ando, 1997). However, prior to our studies, no systematic attempt has been made to examine variations in τ_e for singing voices.

In order to achieve the temporal blending of an operatic singing voice with a given sound field (see APPENDIX A), our pilot study investigated the features of τ_e for the voices of amateur tenors for a single-tone vowel and for musical pieces that were sung in different styles of falsetto (head register) and operatic singing (chest register) (Kato and Ando, 2002). It was too complicated to control τ_e for singing voice signals in the case of musical pieces that included transient musical tones, CV (consonant-vowel sounds), and voice onsets and offsets; hence, we focused on determining a broad variation of $(\tau_e)_{min}$ within the steady-state part of single-tone vowels by alternating between changing such singing styles. Our second study described the τ_e characteristics of a professional operatic singing voice for the steady-state part of vowels in relation to the following factors: (1) vowel selection such as /ah/, /eh/, and /oh/, (2) tone pitch, (3) vibrato rate, (4) vibrato extent, and (5) intonation (Kato et al., 2004a). A negative correlation between $(\tau_e)_{min}$ values and vibrato extent was observed. The aim of the present study is to examine whether singers can consciously vary the value of $(\tau_e)_{min}$ of the steady-state part of vowels in terms of the "subjective degree of vibrato."

2. EXPERIMENTAL METHOD AND ANALYSIS
2.1 Subjective and recording conditions
Ten subjects including (1) professional singers, (2) students of a conservatory, and (3) trained amateur singers participated in this experiment. The subjects' profiles are listed in Table 1. The voices of the singers were recorded in an anechoic chamber using a 1/2-inch condenser microphone placed at a distance of 25 cm in front and 5 cm to the side of the mouth. The signals were sampled at 44.1 kHz.

2.2. Recordings Tasks
The focus areas of the present study are the three musical factors that are instructed to and interpreted by the singers for their performance. Although vowel selection and tone pitch are not the main focus of the present study, these two factors are also included, because our previous study with a small data set observed a significant contribution of both factors to the $(\tau_e)_{min}$ values (Kato et al., 2004a). The subjects were asked to sing five different vowels (/ah/, /eh/, /ih/, /oh/, and /uh/) six

Table 1. Status of the ten singers.

Singer	Voice classification	Professional education (years)	Musical experience (years)	Age (years)
Sop. 1		27	33	43
Sop. 2	Soprano	4	20	25
Sop. 3		4	17	22
Mez. 1		26	26	42
Mez. 2	Mezzo Soprano	14	30	34
Mez. 3		4	14	22
Ten. 1		0	20	45
Ten. 2	Tenor	0	20	25
Bar.	Baritone	39	39	54
Bas.	Bass	27	27	38

times using three different pitches. The following tone pitches were set for each voice classification: Soprano (F4, C5, F5), Mezzo soprano (C4#, G4#, C5#), Tenor (F3, C4, F4), Baritone (C3#, G3#, C4#), and Bass (B2b, F3, B3b). The target frequencies were given by a pitch-pipe before changing the tone pitch. The duration of each tone was set at 1.8 s, and it was maintained by the aid of a visual metronome located 1.5 m in front of the singer.

Since one of the main purposes of this study is to investigate the effect of the "subjective degree of vibrato," this was explained to the singers in the following manner:

* Minimum-vibrato: The singer was asked to sing the smallest degree of vibrato used in musical pieces.

* Medium-vibrato: The singer was asked to sing without any special conscious effort to add extra vibrato.

* Maximum-vibrato: The singer was asked to sing the highest degree of vibrato used in musical pieces.

"Subjective singing volume" may also be a musical factor that influenced the $(\tau_e)_{min}$ values (Kato et al., 2005); however, in this study, it is maintained constant at mezzo-forte (mf) in order to focus on not the extreme but the normal singing volume production interpreted by each individual singer.

Approximately one hour was required to obtain a recording of all the samples from each subject. During this recording session, none of the singers reported any fatigue effects. The total material comprised 2700 tones (5 vowels × 6 trials × 3 pitches × 3 subjective degrees of vibrato × 10 subjects).

2.3. Analysis of the effective duration (τ_e) of the runnig autocorrelation function

In order to statistically isolate the steady-state part of a vocal segment, the segments at the beginning and tail of the tone were eliminated in advance, by using the following rules:

(Rule I) Segments with A-weighted sound pressure level (SPL) should be over 60 dBA, and
(Rule II) Segments with fundamental frequency (F_0) should be ±150 cents (1.5 semi-tones) from the mean fundamental frequency of the tone.

The tone duration for the remaining segment was 1.9 ± 0.2 (s.d.) [s].

The r-ACF as a function of time lag (τ) can be formulated as:

$$\phi_p(\tau) = \phi_p(\tau; t, T) = \frac{\Phi_p(\tau; t, T)}{\left[\Phi_p(0; t, T)\Phi_p(0; \tau + t, T)\right]^{1/2}} \quad (1),$$

where

$$\Phi_p(\tau; t, T) = \frac{1}{2T} \int_{t-T}^{t+T} p'(s)p'(s+\tau)ds \quad (2),$$

In the above expression, 2T represents the integration interval and $p'(s) = p(t)*s(t)$. The function $p(t)$ denotes the amplitude of the original waveform of the recorded signals, and the function $s(t)$ was chosen as the impulse response of the A-weighting filter corresponding to ear sensitivity. Note that the r-ACF is normalized by the geometric mean of the energy at t and energy at $\tau + t$ and should not be normalized by only the energy at t; this ensures that the normalized r-ACF satisfies $\phi_p(0) = 1$ and $\phi_p(\tau) \le 1$ at $\tau > 0$. The procedure for calculating the r-ACF is shown in APPENDIX B. If the cutout segment of the signals are processed by applying an FFT algorithm (FFT method A, see APPENDIX B) without the concept of time-lag, window functions such as Hamming, Hanning, or Blackman are sometimes used, for example, for technically detecting the fundamental frequency (F_0) of a sound signal. However, for simplification, we used a rectangular window as a time window function for the sound data in this study. The shape of the window function affecting the decay rate of the r-ACF should be chosen carefully and cannot be easily be determined at the present stage.

Figure 1 (a) illustrates an example of the logarithm of the absolute value of the r-ACF as a function of the time lag τ. Note that this figure does not show a decrease in the sound pressure level of the room sound field, such as a reverberation curve; however, the figure represents the decrease in the absolute value of the signal autocorrelation. The effective duration (τ_e) is defined by the time lag at which the envelope of the r-ACF becomes −10 dB, which can be obtained from extrapolating the decay rate to −10 dB (i.e., $|\phi_p(t)| = 0.1$). In the previous studies (Ando et al., 1989; Ando, 1998; Ando et al., 2000; Ando, 2002; Kuroki et al., 2004; Nakayama, 1984; Noson et al., 2000; Noson et al., 2002; Sato et al., 2000), the envelope decay for the "initial part" of the logarithm of the r-ACF of musical pieces have been considered to be linear. In this study, a straight line

τ_e

5 dB

(a)

:Peak detect (This study)
:Peak detect (250 ms)
:Backward Integ. (50 ms)
:Backward Integ. (250 ms)

(b)

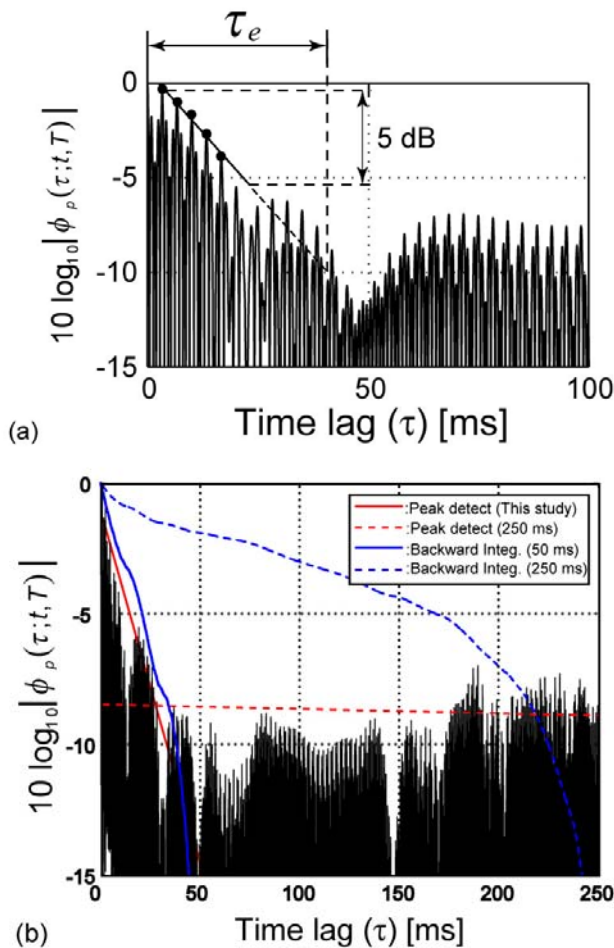Fig. 1. Example of the determination of the effective duration (τ_e) of the r-ACF of an operatic-style sung vowel: (a) peak detecting method used in this study, (b) comparison of the peak detecting method and the backward integration method.

obtained by using the least mean squares method was fitted to the major local peaks corresponding to multiples of the fundamental period, in the range from the amplitude of the first major peak to the amplitude of first major peak subtracted by 5 dB. However, if the entire major local peaks from the origin to $\tau = 50$ ms exceed the amplitude of first major peak subtracted by 5 dB, then a straight line was fitted to all the major local peaks (see APPENDIX C). It may be possible to determine the envelope of the ACFs by using the Schroeder Backward Integration method (Schroeder, 1987), which is applied for calculating the reverberation times from the decay curves of room impulse responses (Fig. 1 (b)); however, we applied the former "peak detecting method" in order to compare the present result with the results obtained from the previous studies.

It has been recommended that the value of the integration interval (2T) of the r-ACF should be around 30 times the typi-

cal value for $(\tau_e)_{min}$ (Mouri et al., 2001). However, our previous study showed that the choice of 2T had little effect on the value of $(\tau_e)_{min}$, as far as the operatic singing voice of vowels is concerned (Kato et al., 2004a; Kato et al., 2004b). For this study, we selected fixed integration times for all data at 2T = 500 ms. Figure 2 provides an example of the running τ_e of an operatic singing voice signal as a function of time.

2.4. Analysis of variance (ANOVA) and linear prediction models of $(\tau_e)_{min}$

The purpose of this study is to examine whether the singers can consciously vary the $(\tau_e)_{min}$ values using musical cues without having a knowledge of the physical parameters such as vibrato rate, vibrato extent, jitter, and shimmer. Hence, the following three musical factors that were determined by our instruction to the singers as described in subsection 2.2 were assumed to be qualitative explanatory variables for $(\tau_e)_{min}$ in this study: (1) subjective degree of vibrato (3 levels: minimum-vibrato, medium-vibrato, and maximum-vibrato), (2) tone pitch (3 levels: with absolute height depending on vocal part), and (3) vowel selection (5 levels: /ah/, /eh/, /ih/, /oh/, and /uh/). For the case when individual singer variations are not necessary, the factor for "Subject" (10 levels) can be treated as a residual; however, this study included this factor among the explanatory variables because of the focus on individual singer variations rather than an average across all the singers. The results of analysis of variance (ANOVA) including the three musical factors and the factor for "Subject" might reflect the relative importance of the musical factors versus the individual singer variations. Consequently, a four-way ANOVA was performed.

The efficiency of the linear prediction model of the $(\tau_e)_{min}$ values by the three musical factors and the factor for "Subject" is also of interest of us, because such linear prediction
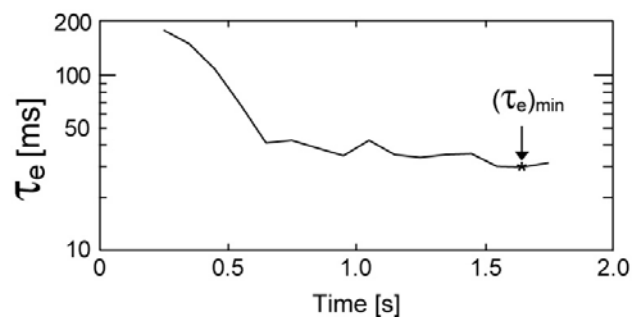

Fig. 2. Example of the measured τ_e values of the r-ACF with a 100-ms stepping interval (2T = 500 ms) of a recording of a G3# "oh" vowel, as sung by a professional baritone singer at mf (mezzo forte).

J. Temporal Des. Arch. Environ. 6(1), December 2006

Kato et al. 38

models, which can be used for the linear variation models of the $(\tau_e)_{min}$ values are convenient for the singers. In this study, the linear prediction process was divided into three parts. Firstly, we performed a linear prediction by using only the three musical factors (the simplest case). Second, we performed a linear prediction by using the three musical factors and the factor for "Subject". The obtained result will reflect the effectiveness of the relative variation of $(\tau_e)_{min}$ and not the absolute. Lastly, we performed a linear prediction by using the three musical factors for each individual singer. The result will make it possible to compare the variation of $(\tau_e)_{min}$ by each individual singer among the ten singers.

3. RESULTS

Figure 3 shows the distribution of the $(\tau_e)_{min}$ values for sung vowels (10 singers). The $(\tau_e)_{min}$ values show an approximately normal distribution on a logarithmic scale and not a linear scale. The $(\tau_e)_{min}$ values for individual tones ranged between 5.5 ms and 750 ms, and the geometric mean of the individual singers was ranged between 21 ms (Sop. 2) and 90 ms (Mez. 3). The geometric mean across all singers was 36 ms.

Table 2 lists the results of ANOVA after fitting a model with four variables to the data that included all 2700 sung vowels. We performed the full factorial ANOVA, in which the main effects of each variable and the interactions between two variables were included in the model. The measured total contribution ratio for the model was 80.4%.

For practical convenience, the sum of the squares of each main effect and interaction were normalized by the total of the sum of the squares; these are shown along with the main contributing factors in a bold font in Table 2(a). The fourth column of Table 2(b) lists each contribution ratio to the value of $\log_{10}(\tau_e)_{min}$ in the percentile scale. Since the three musical
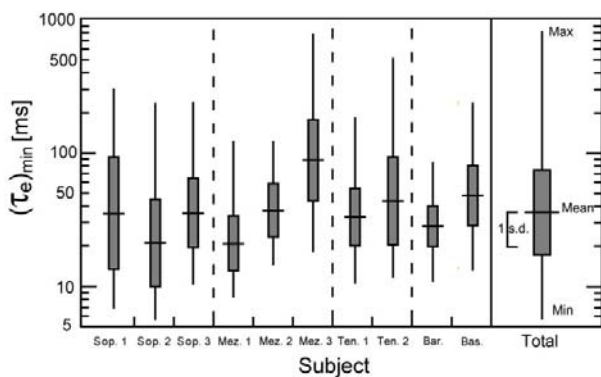

Fig. 3. Measured distribution of the $(\tau_e)_{min}$ values of sung vowels for ten singers.

Table 2. Results of the analysis of variance (ANOVA).

(a)

Factor	DF	Sum of square	Mean square	F-ratio
Model	449	259	0.58	42.29
Error	2250	31	0.01	
Total	2699	**289**		$p < 0.001$

(b)

Factor	DF	Sum of square	Contribution ratio [%]
Vib.	2	78.2	**27.0**[**]
TP	2	0.7	0.2[**]
Vowel	4	12.7	4.4[**]
Subject	9	82.7	**28.6**[**]
Vib.*TP	4	1.1	0.6[**]
Vib.*Vowel	8	0.6	0.2[**]
TP*Vowel	8	2.5	0.8[**]
Vib.*Subject	18	20.0	6.9[**]
TP*Subject	18	22.7	7.9[**]
Vowel*Subject	36	11.8	4.1[**]
Total		259	**80.4**

Vib. = Subjective degree of vibrato, TP = Tone pitch. The values larger than 10% are marked in bold. ** 1% significant level.

factors–subjective degree of vibrato (Vib.), tone pitch (TP), and vowel selection (Vowel)–are dependant upon musical expression, they can be indicated on a musical score (TP and Vowel). Further, they are also influenced by individual singer interpretation styles (Vib., TP, and Vowel).

In order to predict the absolute value of $(\tau_e)_{min}$, a linear prediction model that employs the main effects of the three factors and excludes the factor of "Subject" can be formulated as

$$\log_{10}(\tau_e)_{min} \approx a_1(\text{Vib.}) + a_2(\text{Pitch}) + a_3(\text{Vowel}) + c \qquad (3)$$

where a_1(Vib.), a_2(Pitch), a_3(Vowel), and c are the values calculated by a multiple regression analysis with dummy variables; these values were fitted to the data. This statistical analysis is mathematically equivalent to a multiple-dimensional-factor analysis (Hayashi, 1952; Hayashi, 1954a; Hayashi, 1954b) (see APPENDIX D). Table 3 lists the values of a_1(Vib.), a_2(Pitch), and a_3(Vowel), which are termed as "category score" in Hayashi's model. The value of the constant c in Equation (3) is equal to the logarithm of the geometric mean of $(\tau_e)_{min}$ across the singers (see Fig. 3: $c = \log_{10}(36)$). The summation of the contribution ratios for these three variables–where Vib., TP, and Vowel = 27.0%, 0.2%, and 4.4%, respectively–was 31.6%.

J. Temporal Des. Arch. Environ. 6(1), December 2006

Kato et al. 39

Although the (Vib.) had a relatively high contribution, this model did not result in good predictions for the absolute values of $(\tau_e)_{min}$ (Fig. 4 and Table 2(b)).

The contribution ratio of the main effect of "Subject" was high (28.6%). When this factor is added to the above mentioned model expressed in Equation (3), we can formulate the following:

$$\log_{10}(\tau_e)_{min} \approx a_1(Vib.) + a_2(Pitch) + a_3(Vowel) + a_4(Subject) + c'$$
(4)

where $a_1(Vib.)$, $a_2(Pitch)$, and $a_3(Vowel)$ have identical values to those in Equation (3) and listed in Table 3. The expression "$a_4(Subject) + c$" denotes the logarithm of the geometric mean value of $(\tau_e)_{min}$ for each singer (see Fig. 3; the values range between $\log_{10}(21)$ (Sop.2) and $\log_{10}(90)$ (Mez. 3)). After fitting this model, the total contribution ratio became 60.2% as shown in Fig. 5 and Table 2(a).

The effects of interaction between "Subject" and the other three factors were rather high (Table 2(b)). We decided to further examine the data by comparing linear prediction models for individual singers, which were formulated as follows:

$$[\log_{10}(\tau_e)_{min}]_{(j)} \approx b_{1(j)}(Vib.) + b_{2(j)}(Pitch) + b_{3(j)}(Vowel) + c_{(j)},$$
$$j = 1, 2, ..., 10$$
(5)

where $b_{1(j)}(Vib.)$, $b_{2(j)}(Pitch)$, and $b_{3(j)}(Vowel)$ are calculated by

Table 3. Values of $a_1(Vib.)$, $a_2(Pitch)$, and $a_3(Vowel)$ in Equation (3).

Factor (Item)	(Category)	Coefficient (Category score)
	Minimum-vibrato	+0.21
$a_1(Vib.)$	Medium-vibrato	+0.01
	Maximum-vibrato	−0.21
	Low	+0.02
$a_2(Pitch)$	Middle	−0.01
	High	−0.02
	/ah/	−0.08
	/eh/	−0.04
$a_3(Vowel)$	/ih/	−0.03
	/oh/	+0.03
	/uh/	+0.12

the multiple regression analyses with dummy variables for the individual singers; these values were fitted to the data. Figure 6 shows the values of $b_{1(j)}(Vib.)$, $b_{2(j)}(Pitch)$, and $b_{3(j)}(Vowel)$. The $c_{(j)}$ values denote the logarithm of the geometric mean value of $(\tau_e)_{min}$ for each singer and are identical to "$a_4(Subject) + c$" mentioned in the preceding paragraph. Although these values greatly varied among the individual singers, Vib. showed a similar trend for the individual singers, in which all the $(\tau_e)_{min}$ values decreased with a greater subjective degree of vibrato. Table 4 lists the contribution ratios of Vib., TP, and Vowel to the $(\tau_e)_{min}$ values for each singer. As noted, the values varied greatly among the singers. The total contribution
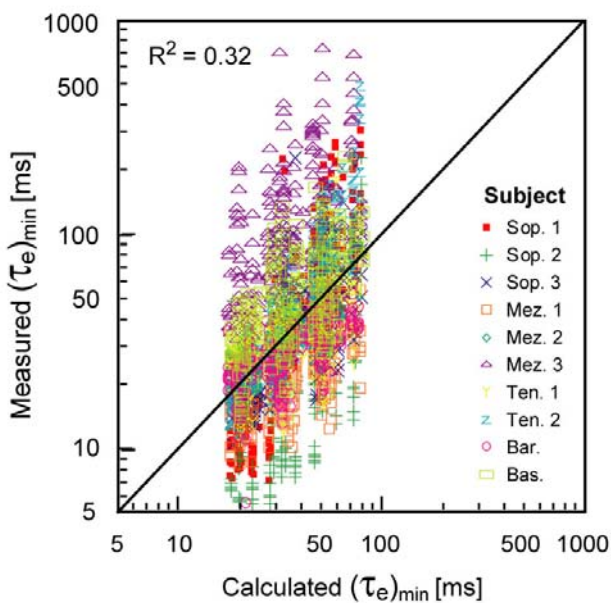

Fig. 4. Relationship between the measured $(\tau_e)_{min}$ values and those calculated by fitting models formulated using Equation (3).
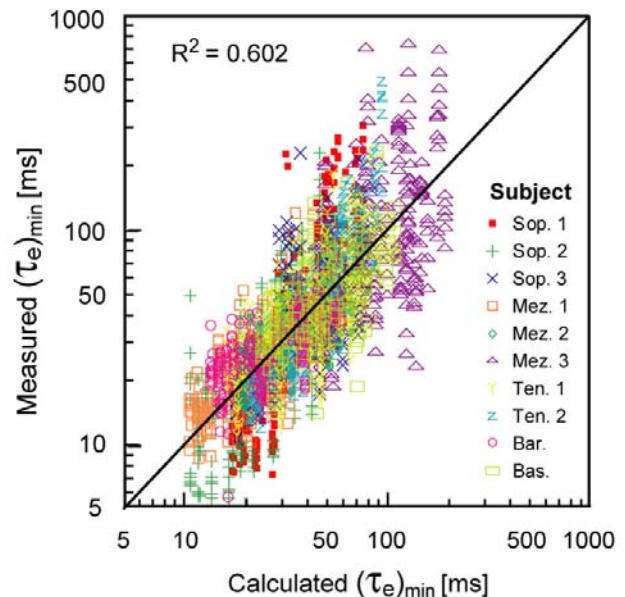

Fig. 5. Relationship between the measured $(\tau_e)_{min}$ values and those calculated by fitting models formulated using Equation (4).

J. Temporal Des. Arch. Environ. 6(1), December 2006
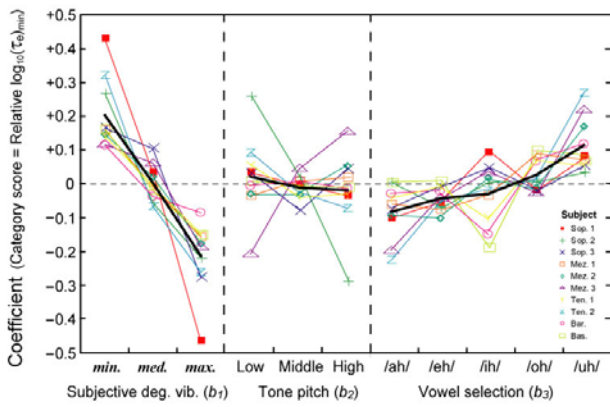
Kato et al. 40

Fig. 6. Coefficients obtained from multiple regression analyses with dummy variables for ten singers. : Mean values of coefficients among ten singers (corresponding to a_1(Vib.), a_2(Pitch), and a_3(Vowel) in Equations (3) and (4) and listed in Table 3).
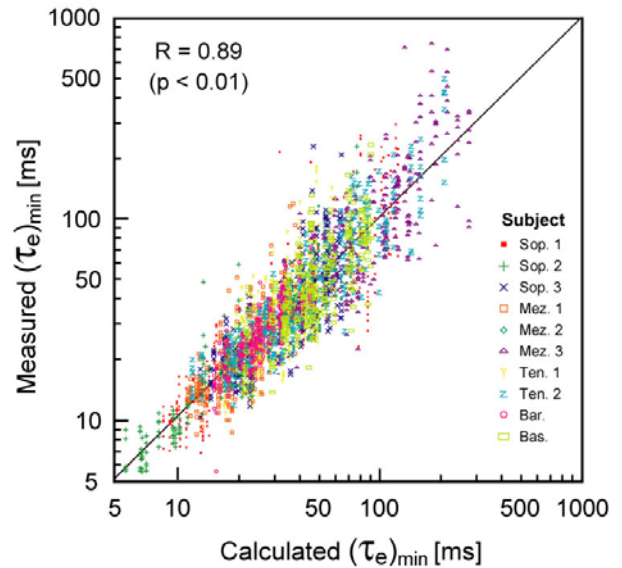

Fig. 7. Relationship between the measured $(\tau_e)_{min}$ values and those calculated by fitting models formulated using Equation (5) for individual singers.

ratio for each singer's model varied between 42% and 86%. The contribution ratios of vowel selection to the value of $(\tau_e)_{min}$ for the three sopranos (1%−3%) were smaller than those for the other singers (9%−36%). Figure 7 shows the relationship between the measured $(\tau_e)_{min}$ values and those calculated by Equation (5) for each singer. Although the residuals with respect to (1) interactions of factors and (2) intra-individual changes within 6 trials were included, the correlation coeffi-

cient between the measured and calculated $(\tau_e)_{min}$ values was sufficiently high to reach 0.89 ($p < 0.01$).

4. DISCUSSION

Operatic-style singing voices for ten trained singers, performing variably pitched vowels at a mezzo-forte (mf) singing volume, were recorded and analyzed (1) to determine the possible existing range of $(\tau_e)_{min}$ of variably pitched vowels with variation in the subjective degree of vibrato and (2) to examine whether operatic singers can consciously vary the $(\tau_e)_{min}$ values by means of the "subjective degree of vibrato."

4.1. Range of $(\tau_e)_{min}$ values for vowels sung by trained operatic singers

Figure 3 shows that the log-normally distributed values of $(\tau_e)_{min}$ ranged between 5.5 ms and 750 ms. Let us suppose that a singer produces nearly nonperiodic $((\tau_e)_{min} \approx 0$ ms$)$ or nearly perfect periodic signals $((\tau_e)_{min} \approx \infty)$ with extreme values of $(\tau_e)_{min}$, e.g., by the combination of a softer voice or louder voice (Kato et al., 2005). Even in such a case, it is interesting that the present result for the mezzo-forte (mf) singing volume is broader than the entire variation of previously reported values for continuous speech signals $((\tau_e)_{min} \approx 10$ ms$)$ and musical pieces $((\tau_e)_{min} \approx 20\text{-}200$ ms$)$ (Ando 1998). The measured $(\tau_e)_{min}$ values for ten musical pieces that included solo instrument and orchestra performances ranged between 19 ms (violin solo) and 187 ms (piano, classical mood) in a previous report (Kuroki et al., 2004). A similar broad range was also

Table 4. Contribution ratios of the subjective degree of vibrato, tone pitch, and vowel to the $(\tau_e)_{min}$ values for each individual singers. The maximum and minimum for each variable are marked in bold.

Singer	Contribution ratio [%]			
	Subjective degree of vibrato	Tone pitch	Vowel	Total
Sop. 1	**77**[**]	<1[*]	3[**]	81[**]
Sop. 2	38[**]	**47**[**]	**1**[**]	**86**[**]
Sop. 3	57[**]	5[**]	3[**]	65[**]
Mez. 1	41[**]	1[*]	11[**]	53[**]
Mez. 2	53[**]	5[**]	29[**]	86[**]
Mez. 3	**18**[**]	24[**]	18[**]	60[**]
Ten. 1	30[**]	4[**]	9[**]	**42**[**]
Ten. 2	55[**]	5[**]	24[**]	83[**]
Bar.	32[**]	<1	**36**[**]	68[**]
Bas.	33[**]	<1	20[**]	53[**]

* 5 % significant level. ** 1 % significant level.

observed in our previous investigation of professional operatic singing voices for vowels sung without varying the subjective degree of vibrato (Kato et al., 2004a). In this study, the upper range of $(\tau_e)_{min}$ is higher than that in the previous study. The present study suggests that (1) a conscious vocal effort to reduce the subjective degree of vibrato results in more periodic characteristics and larger $(\tau_e)_{min}$ values (Fig. 6) and (2) there are individual singers who naturally sing with larger $(\tau_e)_{min}$ values, e.g., Mez. 3 (Fig. 3).

4.2. Potential for the variation of $(\tau_e)_{min}$ values

The ability of the singers to vary the $(\tau_e)_{min}$ values in terms of (1) subjective degree of vibrato, (2) tone pitch, and (3) vowel selection was examined. The results for a number of subjects that were formulated on the basis of Equation (3) show that the predictability of the absolute value of $(\tau_e)_{min}$ by the variables Vib., TP, and Vowel is limited (Fig. 4 and Table 2). Thus, it may be difficult to develop a unique model for blending the operatic singing voice with an opera house at the present stage which can be useful for all the singers.

It is interesting to note that the contribution ratio of the main effect of "Subject" was high (Table 2(b)). This may be due to wide variation in the voice quality and phonation style (Sundberg, 1987; Titze, 1994) as well as vibrato extent (Prame, 1997) of individual singers. A negative correlation between vibrato extent and $(\tau_e)_{min}$ values was observed in our previous study (Kato et al., 2004a). The present results suggest that intra-individual variability of $(\tau_e)_{min}$ in terms of Vib., TP, and Vowel, is relatively more realizable; this is because the common model for all singers that was formulated using Equation (4) resulted in a total contribution ratio of 60.2% (Figure 5 and Table 2). Thus, for blending the operatic singing voice with a given opera house, it may be more effective for singers to be aware of their individual averaged values of $(\tau_e)_{min}$ which can be measured as described in this study.

The results presented in Fig. 6 and Table 4 show that the contribution ratios of Vib., TP, and Vowel varied greatly among the individual singers. Only Vib. had a recognizable predictive trend toward higher relative contribution to the $(\tau_e)_{min}$ values. Therefore, one of the most important factors that influences the subjective responses of both listeners and performers to the sound field of the performing hall, which can be used in the "lifeline of music" (Tsutsumi, 2005), may be the "subjective degree of vibrato." As an application, for example, it might in the effective selection of a music program aimed at reducing the subjective degree of vibrato for performances in a hall where the sound fields might have longer delayed reflections and/or a longer reverberation time.

Note that the greater subjective degree of vibrato causes the $(\tau_e)_{min}$ values to decrease (Fig. 6). The present observed phenomenon might be closely related to changes in the vocal production mechanism when the subjective degree of vibrato is changed. The effort to produce a greater subjective degree of vibrato might cause non-periodic voice signal production from the vocal fold to the mouth. Although the effect of the subjective degree of vibrato on the $(\tau_e)_{min}$ values is difficult to determine in mathematical terms, it would be reasonable to question whether the amplitude of r-ACF might increase again around the vibrato cycle (160–200 ms) (see Fig. C1 (b)). If pure tones are modulated with perfect periodicity, the amplitude of r-ACF increases again to 0 dB at the modulation period. Also, it would be informative if we analyze the $(\tau_e)_{min}$ values for several synthetic voice signals as a function of vibrato rate and vibrato extent. A topic of discussion could be ascertaining the more dominant effect between the effect of vibrato content and that of random variations in phonation period time. In order to describe the effect of the subjective degree of vibrato in relation to the qualitative parameters, the acoustic analyses of signals such as sound level, vibrato rate, vibrato extent, intonation, jitter, and shimmer must be addressed in future studies.

The results obtained from the $(\tau_e)_{min}$ values of the five vowels (/ah/, /eh/, /ih/, /oh/, and /uh/) are noteworthy (Fig. 6). The average $(\tau_e)_{min}$ values for the ten singers was in the following order: /ah/ < /eh/ < /ih/ < /oh/ < /uh/. This finding is useful for discussing the quasi-periodicity of sung vowels in relation to the production mechanism of the singing voice. The present result may be closely related to those of previous studies on vowels such as the position of the tongue or the opening of the mouth and jaw in singing (Titze, 1994), and this physiological aspects of singing may need to be addressed in future studies.

4.3. Individual singer variations

It is interesting to note that the interactions between "Subject" and the other three factors were fairly high (Table 2(b), Fig. 6); the reason for such a high interaction is discussed now. Firstly, with regard to the high interaction of Vib.(Subjective degree of vibrato) with Subject, this may be related to the variations in the vocal production mechanism of individual singers, and it requires more research. It can be hypothesized that the reason for the present observation may be related to typical individual variations in vibrato extent (Prame, 1997) and not vibrato rate (Prame, 1994). In our previous study, a significant negative correlation was observed between the $(\tau_e)_{min}$ values and vibrato extent, but the correla-

J. Temporal Des. Arch. Environ. 6(1), December 2006

Kato et al. 42

tion between the value of $(\tau_e)_{min}$ and vibrato rate was small (Kato et al., 2004a).

Secondly, the high interaction of TP(Tone Pitch) with Subject may be closely related to variations in the changes of the vocal register of an individual singer (Sundberg, 1987; Titze, 1994). The vocal register can be divided into a minimum of three parts in the middle range of one octave for each vocal part (see subsection 2.2 of this manuscript): falsetto (head), mezzo falsetto, and middle (chest). Trained operatic singers tend to use these three types of registers smoothly and individually (Sundberg, 1987; Titze, 1994). In the normal pitch range of the falsetto (head) register, the waveform of the recorded voices is more similar to the waveform of pure tones than that in the case of the middle (chest) register. Our previous study suggested that this factor was responsible for the $(\tau_e)_{min}$ values for the falsetto register becoming larger than that for the middle (chest) register (Kato and Ando, 2002). For example, in this study, Mez. 3 may have been sung with a falsetto (head) register; thus increasing the $(\tau_e)_{min}$ values at a high pitch range, while Sop. 2 may have been sung with a chest register, thus decreasing the value of $(\tau_e)_{min}$ at this range. In order to confirm this speculation, the glottal open quotient, which is strongly related to the vocal register, should be physiologically measured by using the DECOM (DEgg: Correlation-based Open quotient Measurement) method (Henrich et al., 2005), and to be addressed in future studies.

Finally, the high interaction of Vowel with Subject can be expressed as follows. Among the five vowels, the individual singer variation for a given vowel was especially high for /ih/ and /uh/ vowels. This result may be closely related to previous studies on vowels that involved factors such as the position of the tongue or the opening of the mouth and the jaw in singing (Titze, 1994); this requires physiological studies of singing to be addressed as a future study. The fact that the contribution ratios of vowel selection to the $(\tau_e)_{min}$ values for the three sopranos were smaller than those for the other singers (Table 4) is encouraging. This result is as expected since vowels tend to be sung in a more similar manner at high pitches (F_0–F_1 matching).

The present result shows that linear prediction models for individuals formulated in Equation (5) can allow for reasonably accurate prediction as shown in Table 4, Fig. 6, and Fig. 7. Therefore, it is practically effective to find the dominant musical factors which contribute to $(\tau_e)_{min}$ values for each individual singer. This can be achieved by the personal measurements similar to those performed in this study along with the singer's request.

5. CONCLUSIONS

A large data set of the subjectively significant temporal parameter $(\tau_e)_{min}$ of the steady-state part of operatic-style sung vowels with a mezzo forte (mf) singing volume was used: the resulting observations were as follows:

1. Typically, the subjective degree of vibrato has a higher contribution to the value of $(\tau_e)_{min}$, and it decreases with a greater subjective degree of vibrato (see Tables 2 , 3, and 4 and Fig. 6).

2. The selection of a subjective degree of vibrato, tone pitch, and vowel can help achieve a relative variation for the mean values of $(\tau_e)_{min}$ of each individual singer, rather than a determination of the absolute value of $(\tau_e)_{min}$ (compare Fig.4 and Fig. 5).

3. The contribution ratio of the subjective degree of vibrato, tone pitch, and vowel selection to the value of $(\tau_e)_{min}$ depends on the individual singer (see Fig. 6 and Table 4).

.

APPENDIXA:

The role of room acoustics in the effective duration (τ_e) of the running autocorrelation function (r-ACF) of a sound signal With regard to the temporal aspects of room acoustics, reflected sounds with various delays and amplitudes superimpose the direct sound at the ear positions of both listeners and performers. In order to assess the significance of the temporal blending of the source signal with room acoustics, this superposition effect should be described and investigated. The effective duration (τ_e) of the r-ACF is an important parameter for sound signals, and it can reveal the characteristic temporal aspects of room acoustics. The equations expressing the relationship between the $(\tau_e)_{min}$ values of a sound signal and the preferred temporal conditions of sound fields for both listeners and performers are briefly introduced and summarized here. For the simplest case that is illustrated in Fig. A1, when considering the temporal coherence of sound signals in the time domain, the correlation between a direct sound and the delayed reflections is strongly related to the perceptual

J. Temporal Des. Arch. Environ. 6(1), December 2006
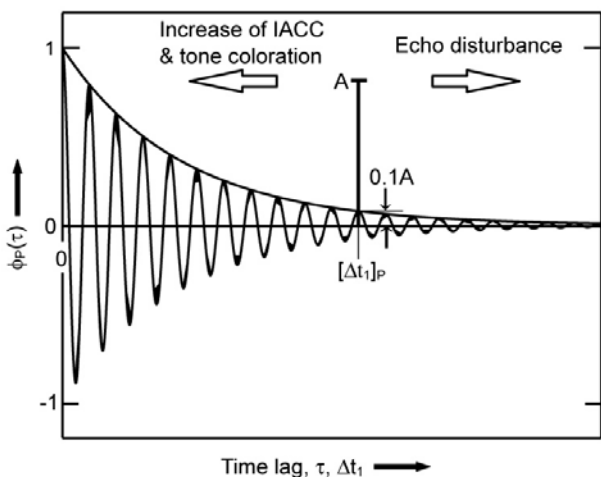
Kato et al. 43

Fig. A1. Subjective attributes of sound fields with a single reflection.

attributes and subjective preferences of the sound fields.

For listeners, using a number of listening tests with musical pieces that have different $(\tau_e)_{min}$ values (Ando et al., 1989; Ando, 1998; Ando et al., 2000; Ando, 2002; Kuroki et al., 2004), the preferred initial time delay ($[\Delta t_1]_p$) between the direct sound and the first reflection from the sound field in a room can be calculated as follows:

$$[\Delta t_1]_p \approx (\log_{10} \frac{1}{k} - c \log_{10} A)(\tau_e)_{min} \qquad (A1)$$

where constants k = 0.1 and c = 1.0, and A is the total amplitude of reflections relative to the direct sound. The preferred subsequent reverberation time $[T_{sub}]_p$ can be calculated as follows:

$$[T_{sub}]_p \approx 23 \, (\tau_e)_{min} \qquad (A2).$$

For performers, the preferred initial time delay ($[\Delta t_1]_p$) between the direct sound and the single reflection from the sound field in a room can be calculated by Equation A1 as well, adjusting constants k and c. The value of A is replaced by the amplitude of the reflection being defined by A = 1 relative to −10 dB of the direct sound as measured at the ear's entrance. This is due to the over-estimation of the reflection by the performer. According to the paired-comparison test with cellists, constants $k \approx 1/2$ and $c \approx 1$ have been obtained (Sato et al., 2000). It is noteworthy that the coefficients k and c for alto-recorder soloists were respectively 2/3 and 1/4 (Nakayama, 1984).

The implications of these three equations can be worded as follows: a non-preferable echo disturbance might occur if the correlation between the direct sound and delayed reflections is smaller than a certain value, then a very long initial time delay Δt_1 causes the echo. Conversely, if the direct sound and delayed reflection is shorter than a certain value, a non-prefer-

able tone coloration may occur due to interference effects between them. The preferred condition may exist between the echo disturbance and the non-preferable tone coloration (see Fig. A1).

APPENDIX B:
Procedure for calculating the running autocorrelation function (r-ACF)

The procedure for calculating the r-ACF, by using the direct method in the time domain, is illustrated in Fig. B1. The ACF is well known as a method for estimating the fundamental frequency (F_0) of a sound signal, as derived by determining the time lag between the origin and the first major peak of the function. Since the fundamental frequency of a musical sound signal is higher than 100 Hz in most cases, the required maximum lag (τ_{max}) to obtain the fundamental period is around 10 ms at most. Yet, in order to obtain the effective duration (τ_e) of the r-ACF, the required length of (τ_{max}) is longer than 50 ms as far as the case of the operatic singing voice of vowels is concerned (see APPENDIX C).

Figure B2 shows a comparison of ACFs and power spectra obtained by different methods. ACFs that employ direct methods are obtained in the time domain. ACFs that use FFT methods, based on the Wiener-Khintchine theorem, are acquired
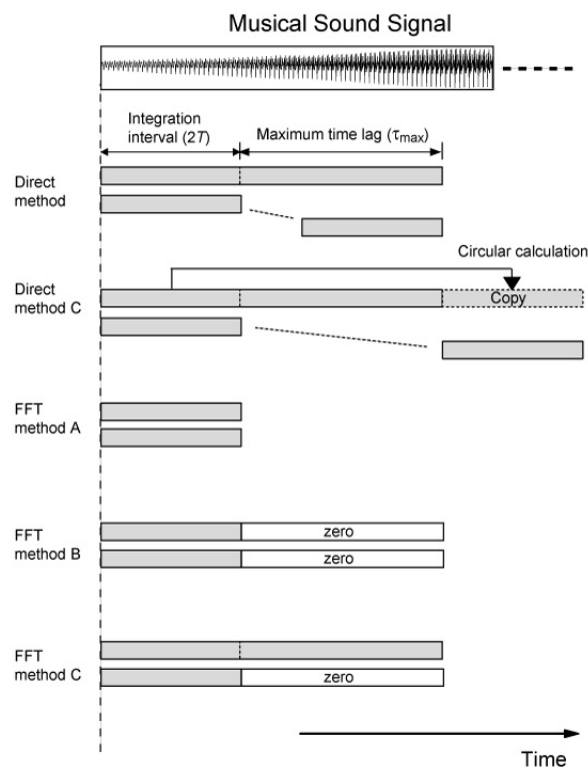

Fig. B1. Direct method to determine the running autocorrelation function (r-ACF) in a time domain.
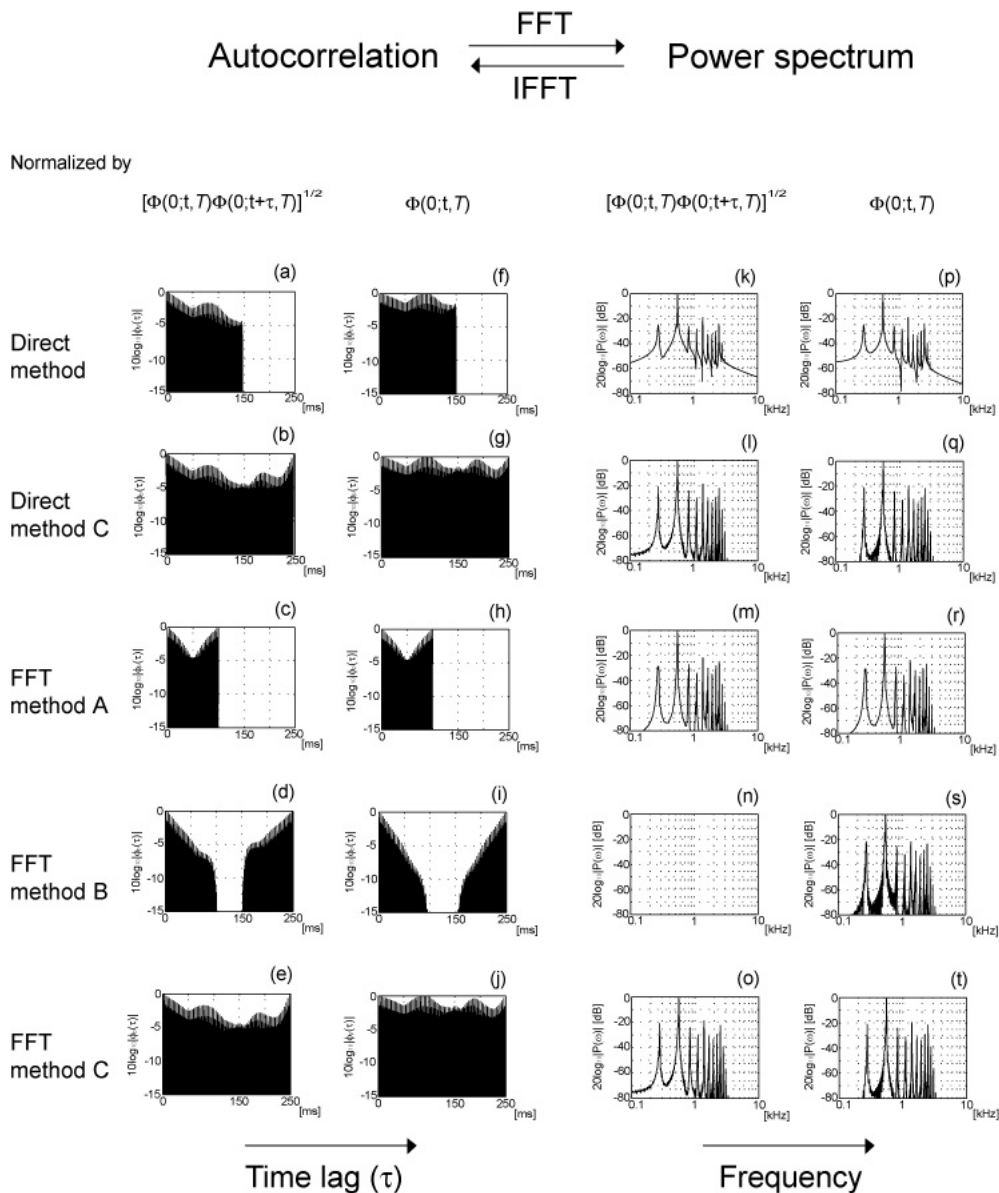
Fig. B2. Comparison of ACFs and power spectra obtained using different methods.

by a transform in the frequency domain by FFT, followed by performing an IFFT calculation. It is important to note that the Wiener-Khintchine theorem is mathematically satisfied only for completely periodic or infinite-length signals, and not mathematically be satisfied for a quasi-periodic signal collected for the analysis of operatic singing voices. Even in a practical situation, a variation in both ACFs and power spectra due to the calculation method is evident (see Fig. B2(a)-(t)). It is not possible to find even one matched pair of the running ACF and running power spectrum for quasi-periodic signals. Thus, we reiterate that the transform methods and their precise definitions should be carefully examined before conducting an analysis of voice signals.

Although "FFT method A" or "FFT method B (method to

avoid circular calculation)" is usually used for the purpose of the fast computation and is accompanied by a window function such as Hamming, Hanning, or Blackman, in order to obtain the ACF corresponding to the direct method, "FFT method C" (see Fig. B2(e)) must be used. If "FFT method C" is chosen instead of the "direct method" for performing a fast calculation, the segment over the maximum time lag is omitted because this segment is obtained from circular calculation, as illustrated in "Direct method C" (see Fig. B2(b) and corresponding Fig. B2(e)).

APPENDIX C:

Definition of "the initial part" of the running autocorrelation function (r-ACF) for calculating its effective duration (τ_e)

The definition of the "initial part" of the r-ACF is briefly described here. Historically (Ando et al., 1989; Ando, 1998; Ando et al., 2000; Ando, 2002; Kuroki et al., 2004; Nakayama, 1984; Noson et al., 2000; Noson et al., 2002; Sato et al., 2000), the envelope decay of the "initial part"of the logarithm of the absolute value of the r-ACF of the source signals of musical sounds has been considered to be linear. However, we need to define the "initial part" of the r-ACF before calculating τ_e because the decay rate of the r-ACF varies, as illustrated in Figure C1. For this study, the "initial part" of the r-ACF was set

Fig. C1. Examples of the r-ACF waveforms and extracted τ_e values of the operatic-style sung vowels.

for both the X-axis (0 ms to 50 ms) and Y-axis (from the amplitude of the first major peak to the amplitude of the first major peak subtracted by 5 dB). This definition of "the initial part" of the r-ACF for calculating τ_e was efficient as far as the operatic singing voice of vowels is concerned.

APPENDIX D:

Method of factor analysis

The method of factor analysis is described here (Hayashi, 1952; Hayashi, 1954a; Hayashi, 1954b). The numeric values to each subcategory of each item are given and synthesized the responses as be concerned with behavior patterns.

In this analysis, all the items do not need to be scalable. The data of cases is used. Here, is an external variable and and are defined as follows: $s = 1, 2,..., R$ (R is the number of items), and $k = 1, 2,..., K$ (K is the number of subcategories in th item). Because each case checks only one subcategory in each item, the behavior pattern of the -case should be synthesized in the form of

$$\alpha_i = \sum_{s=1}^{K_S} X_s(i) = \sum_{s=1}^{R}\{\sum_{k=1}^{K_S}\delta_i(sk)X_{sk}\} \qquad (D1)$$

where

$$\sum_{k=1}^{K_S}\delta_i(sk) = 1 \qquad (D2)$$

and $\delta_i(sk) = 1$, if the -case is under the th subcategory in the th item, and, $\delta_i(sk) = 0$ otherwise. α_i, which is referred to as the total score of the i-case, has a numerical value since X_{sk} has a numerical value.

The correlation coefficient ρ between A and α_i is written as follows:

$$\rho(A,\alpha_i) = \frac{{}^{1}/_{n}\sum_{i=1}^{n}(A_i - \overline{A})(\alpha_i - \overline{\alpha})}{\sigma_A\sigma_\alpha} \qquad (D3)$$

where

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}A_i, \qquad \sigma_A^2 = \frac{1}{n}\sum_{i=1}^{n}(A_i - \overline{A})^2,$$

$$\overline{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\alpha_i, \qquad \sigma_\alpha^2 = \frac{1}{n}\sum_{i=1}^{n}(\alpha_i - \overline{\alpha})^2. \qquad (D4)$$

For the purpose of obtaining a maximum ρ value or to estimate the external variable from the behavior pattern, let $\overline{A} = 0$ and $\overline{\alpha} = 0$ because ρ is invariant under a shift of origin. The score of each subcategory can be determined by solving

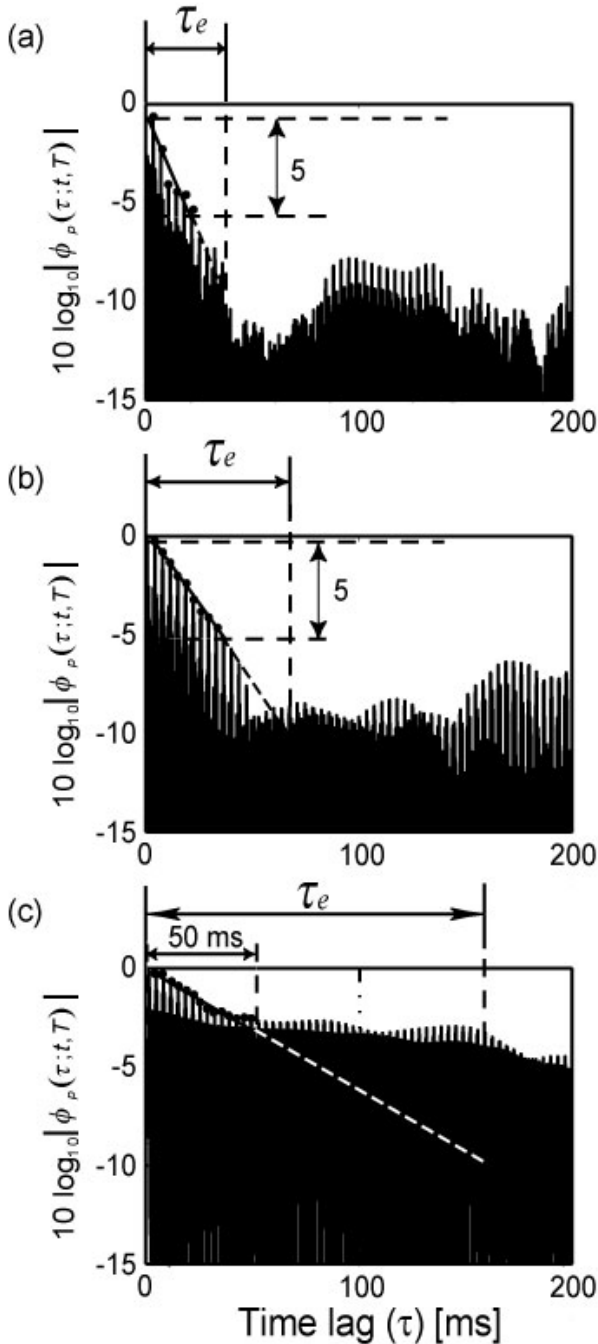$$\frac{\partial\rho}{\partial X_{sk}} = 0 \quad (s = 1, 2,..., R; \ k = 1, 2,..., K_s) \qquad (D5)$$

REFERENCES

Ando, Y., Okano, T., and Takazoe, Y., The running autocorrelation function of different music signals relating to preferred temporal parameters of sound fields, J. Acoust. Soc. Am., Vol. 86, 644-649, 1989.

Ando, Y., Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners. AIP Press/Springer-Verlag, New York, Chapter 3, 4, 6, and 7, 1998.

Ando, Y., Sakai, H., and Sato, S., Formulae describing subjective attributes for sound fields based on the model of auditory-brain system, J. Sound Vib., Vol. 232, 101-127, 2000.

Ando, Y., Sakai, H., and Sato, S., Correlation factors describing primary and spatial sensations of sound fields, J. Sound Vib., Vol. 258, 405-417, 2002.

Arai, T., Kinoshita, K., Hodoshima, N., Kusumoto, A. and Kitamura, T., Effects on suppressing steady-state portions of speech on intelligibility in reverberant environments, Acoust. Soc. & Tech., Vol. 23, 229-232, 2002.

Chiang, W., Chen, S., Huang, C., Subjective assessment of stage acoustics for solo and chamber music performances, Acta Acustica united with Acustica, Vol. 89, 848-856, 2003.

Gade, A.C., Investigation of musician's room acoustic conditions in concert halls. Part I: Method and laboratory experiments, Acustica, Vol. 69, 193-203, 1989a.

Gade, A.C., Investigation of musician's room acoustic conditions in concert halls. Part II: Field experiments and synthesis of results, Acustica, Vol. 69, 249-262, 1989b.

Hayashi, C. On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematico-statistical point of view, Ann. Inst. Statist. Math., III, 69-98, 1952.

Hayashi, C., Multidimensional quantification, I, Proc. Japan Acad., 30, 61-65, 1954a.

Hayashi, C., Multidimensional quantification, I, Proc. Japan Acad., 30, 165-169, 1954b.

Henrich, N., d'Alessandro, C., Doval, B., and Castellengo, M., Glottal open quotient in singing: Measurements and correlation with laryngeal mechanism, vocal intensity, and fundamental frequency, J. Acoust. Soc. Am., Vol. 117, 1417-1430, 2005.

Hodoshima, N., Arai, T., Inoue, T., Kinoshita, K. and Kusumoto, A., Suppressing steady-state portions of speech for improving speech intelligibility in various reverberant environments, Acoust. Soc. & Tech., Vol. 25, 58-60, 2004.

Hodoshima, N., Goto, T., Ohata, N., Inoue, T and Arai, T., The effect of pre-processing approach for improving speech intelligibility in a hall: Comparison between diotic and dichotic listening conditions, Acoust. Soc. & Tech., Vol. 26, 212-214, 2005.

Kato, K. and Ando, Y., A study of the blending of vocal music with the sound field by different singing styles, J. Sound Vib., Vol. 258, 463-472, 2002.

Kato, K., Fujii, K. Kawai, K., Yano, T. and Ando, Y., Blending vocal music with the sound field -the effective duration of autocorrelation function of Western professional singing voices with different vowels and pitches, Proc. International Symposium on Musical Acoustics 2004, Japan, 39-42, 2004a. *Full paper: under review in Acta Acustica united with Acustica.

Kato, K., Fujii, K. Kawai, K., Yano, T. and Ando, Y., Blending vocal music with a given sound field due to the characteristics of the running autocorrelation function of singing voices, J. Acoust. Soc. Am., Vol. 115, 2437, 2004b. *Full paper: under review in Acta Acustica united with Acustica.

Kato, K., Hirawa, T., Kawai, K., Yano, T. and Ando, Y., A method of the blending of operatic singing voices with a given sound field by varying subjective sound volume, Proc. Forum Acusticum Budapest 2005, Hungary, 499-504, 2005. *Full paper: under review in Acta Acustica united with Acustica (invited paper).

Kuroki, S., Hamada, M., Sakai, H., and Ando, Y., Individual preference in relation to the temporal and spatial factors of the sound field: factors affecting individual differences in subjective preference judgments, J. Temporal Des. Arch. Environ., Vol. 4, 29-40, 2004.

Marshall, A. H., Gottlob, D., and Alrutz H., Acoustical conditions preferred for ensemble, J. Acoust. Soc. Am., Vol. 64, 1437-1442, 1978.

Marshall, A. H. and Meyer, J., The directivity and auditory impressions of singers, Acustica, Vol. 58, 130-140, 1985.

Mouri, K., Akiyama, K., and Ando, Y., Preliminary study on recommended time duration of source signals to be analyzed, in relation to its effective duration of the auto-correlation function, J. Sound Vib., Vol. 241, 87-95, 2001.

Nakayama, I., Preferred delay time of a single reflection for performer, Acustica, Vol. 54, 217-221, 1984.

Noson, D., Sato, S., Sakai, H., and Ando, Y., Singer responses to sound fields with a simulated sound field, J. Sound Vib., Vol. 232, 39-51, 2000.

Noson, D., Sato, S., Sakai, H., and Ando, Y., Melisma singing and preferred stage acoustics for singers, J. Sound Vib., Vol. 258, 473-485, 2002.

Otondo, F., Rindel, J. H., The influence of the directivity of musical instruments in a room, Acta Acustica united with Acustica, Vol. 90, 1178-1184, 2004.

Prame, E., Measurements of the vibrato rate of ten singers, J. Acoust. Soc. Am., Vol. 96, 1979-1983, 1994.

Prame, E., Vibrato extent and intonation in professional Western lyric singing, J. Acoust. Soc. Am., Vol. 102, 616-621, 1997.

Sakai, H., Sato, S. and Prodi, N., Orthogonal factors for the stage and pit inside a historical opera house, Acta Acustica united with Acustica, Vol. 90, 319-334, 2004.

Sato, S., Ando, Y., and Ota, S., Subjective preference of cellists for the delay time of a single reflection in performance, J. Sound Vib., Vol. 232, 27-37, 2000.

Schroeder, M. R., New method of measuring reverberation time, J. Audio. Eng. Soc., Vol. 35, 299-305, 1987.

Sundberg, J., The science of the singing voice. Dekalb: Northern Illinois University Press, 1987.

Taguti, T. and Ando, Y., Characteristics of the short-term autocorrelation function of sound signals in piano performances, in Music and Concert Hall Acoustics., Eds., Ando, Y., and Noson, D., Chapter 23, Academic Press, London, 1997.

Ternström, S., Long-time average spectrum characteristics of different choirs in different rooms, STL-QPSR, Vol.3, 14-30, 1989.

Ternström, S., Preferred self-to-other ratios in choir singing, J. Acoust. Soc. Am., Vol. 105, 3563-3574, 1999.

Ternström, S., Cabrera, D., and Davis, P., Self-to-other ratios measured in an opera chorus in performance, J. Acoust. Soc. Am., Vol. 118, 3903-3911, 2005.

Tsutsumi, T., Special lecture, Proc. of the 2nd International Symposium on Temporal. Design, 135, 2005.

Titze, I., Principles of voice production. Prentice-Hall, 1994.

J. Temporal Des. Arch. Environ. 6(1), December 2006

Kato et al. 47

Ueno, K. and Tachibana, H., Experimental study on the evaluation of stage acoustics by musicians using 6-channel sound simulation system, Acoust. Soc. & Tech., Vol. 24, 130-138, 2003.

Ueno, K. and Tachibana, H., Cognitive modeling of musician's perception in concert halls, Acoust. Soc. & Tech., Vol. 26, 156-161, 2005.

Ueno, K., Kanamori, T. and Tachibana, H., Experimental study on stage acoustics for ensemble performance in chamber music, Acoust. Soc. & Tech., Vol. 26, 345-352, 2005.