BLENDING VOCAL MUSIC WITH SOUND FIELDS BY VARIATIONS IN LYRICS AND MUSIC TEMPO

43.75.RS; 43.55.FW; 43.55. HY

Kosuke KATO*; Dennis NOSON**; Yoichi ANDO*

* Graduate School of Science and Technology, Kobe University.

Rokkodai, Nada, Kobe 657-8501

Japan

Tel/Fax: +81-78-803-6471

Email: kkato@kobe-u.com

** BRC Acoustics Inc.

15th Ave. W, Seattle, WA 98119.

U.S.A.

Tel: +1-206-270-8910

Fax: +1-206-270-8690

Email: dnoson@brcacoustics.com

ABSTRACT The initial time-delay Δt_1 between the direct sound and the first reflection and the subsequent reverberation time T_{sub} are usually fixed within a given space. Thus, Concert halls cannot have ideal conditions for all music programs. It has been shown that the most preferred conditions for both listener and performer are determined by the minimum value of the effective duration of the running autocorrelation function (ACF) of sound signals, $(\tau_e)_{min}$ [Y. Ando 1998 Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners. AIP/Springer-Verlag, New York]. An attempt is made here to estimate $(\tau_e)_{min}$ of vocal music by rating various kinds of interpretation styles of singers. The present results showed that $(\tau_e)_{min}$ of the ACF of a voice source varies with lyrics and fluctuation of pitch but not music tempo. Significant findings are: 1) Values of $(\tau_e)_{min}$ are relatively longer in order of: humming > singing with “la” syllables > singing with lyrics > non-voiced (breath noise) singing ($p < 0.01$); 2) Values of $(\tau_e)_{min}$ of fast vocal music are not shorter than those of slow tempo music ($p < 0.05$); 3) Values of $(\tau_e)_{min}$ of vocal music with pitch fluctuation are shorter than those of music with constant pitch ($p < 0.05$).

I. INTRODUCTION

Opera houses and concert halls have unique characteristics, so that no hall is ideally suited to all music styles and all musical instruments. Significant effort has been expended in the last 50 years to identify the acoustical characteristics of sound fields for the audience which enhance the listening experience for any given musical program. More recently, increased attention has been paid to the needs of performers and the nature of the musical sound source, itself, which can influence listener and performer preference for one sound field over another.

The decay characteristic of the autocorrelation function (ACF) of a sound source has been shown to be useful in an analysis of the subjective preference of sound fields [1]. The minimum value of the effective duration of running ACF, $(\tau_e)_{min}$, is a temporal parameter of a given sound source that is closely related to the most preferred temporal conditions for both listeners [1,2] and performers [3-5] (see the APPENDIX). For example, Taguti has shown that tempo, articulation, and damper pedaling are the main elements for determining τ_e of piano music [6]. However, no attempt at examining the τ_e for the singing voice and changes due to singing style has been made.

Studies of the singing voice have been mainly focused on the mechanism of voice source production [7, 8] and power spectrum of the voice source [8, 9]. A number of studies have dealt with the preferred sound field for vocal performers [5, 10-12]. Singer preference studies that consider temporal parameters of the sound source are rare. Only one of these previous singer studies reported the preferred time delay in relation to the temporal parameter, $(\tau_e)_{min}$, of the voice source signal [5]. Our previous study dealt with the singing styles of *falsetto*, *medium falsetto*, and *operatic singing* [13]. Using simple motifs, with five solo singers (tenor), revealed that the style of singing affects the value of $(\tau_e)_{min}$. Values of $(\tau_e)_{min}$ for voice sources singing *falsetto* or *medium falsetto* are much longer than those for *operatic singing*.

Many variations to the singing voice are possible: changes in dynamics, such as crescendo and decrescendo, various kinds of articulation, such as staccato, legato (or tenuto), marcato, sforzando, etc., and others such as vibrato or tremolo. In this study, we have sought to investigate the likely potential for larger variation in the value of $(\tau_e)_{min}$ due to changing the motif's lyrics, stretching note value, and adjusting tempo.

II. EXPERIMENT WITH DIFFERING LYRICS

PROCEDURES

Unlike other musical instruments, there are lyrics for vocal music. In order to examine the difference in $(\tau_e)_{min}$ when the motif were sung in different lyrics, an analysis of $(\tau_e)_{min}$ of the sound source with four singers (1 soprano, 1 alto, 1 tenor, and 1 bass) for a music motif was conducted.

Fig. 1 shows the two musical scores used in this experiment. Motif I (FP) consists of eighth notes and longer, arranged in horizontal intervals of major and minor seconds, with no indication of expression, such as staccato, stress accent, etc. The motif I (FP) was arranged to the motif I (CP) featuring no horizontal interval (constant pitch). The singers were asked to perform motif I (FP) with four different lyrics: humming singing, singing with "la" syllables, singing with real lyrics, and non-voiced (breath noise) singing, similar to whispering.

For the ease of performing, the range of notes was adjusted for each singer. The soprano singer start note for singing was D5, the alto singer G4, tenor singer D4, and bass singer G3. The music tempo was ♩ = 120, and was maintained by the aid of a visual metronome located at a 1.0 m distance in front of the singer. The singer's voice was received by a microphone located 25 ± 1 cm in front of the singer and 5 ± 0.5 cm to the side of the singer's mouth. Singers performed in an anechoic room.

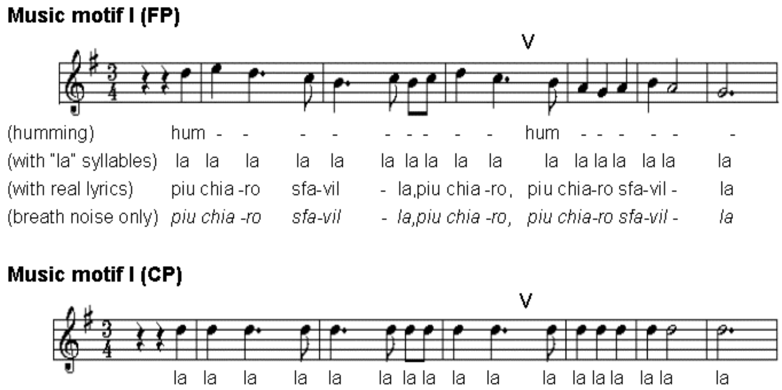

Fig. 1 Music scores of the two motifs used in the experiment.

RESULTS

Fig. 2 shows examples of measured τ_e values of the running ACF for four different lyrics with a 100-ms interval as a function of time. As the recommended integration interval $(2T)_r$ is around $30(\tau_e)_{min}$ [14], $2T$ = 1.0s was selected for each analysis. The mean values and standard deviations of $(\tau_e)_{min}$ obtained in ten trials for each subjects are listed in Table 1. Analyzing the variance showed that values of $(\tau_e)_{min}$ for each subject is relatively longer as follows: humming > singing with "la" syllables > singing with lyrics > non-voiced singing ($p < 0.01$).

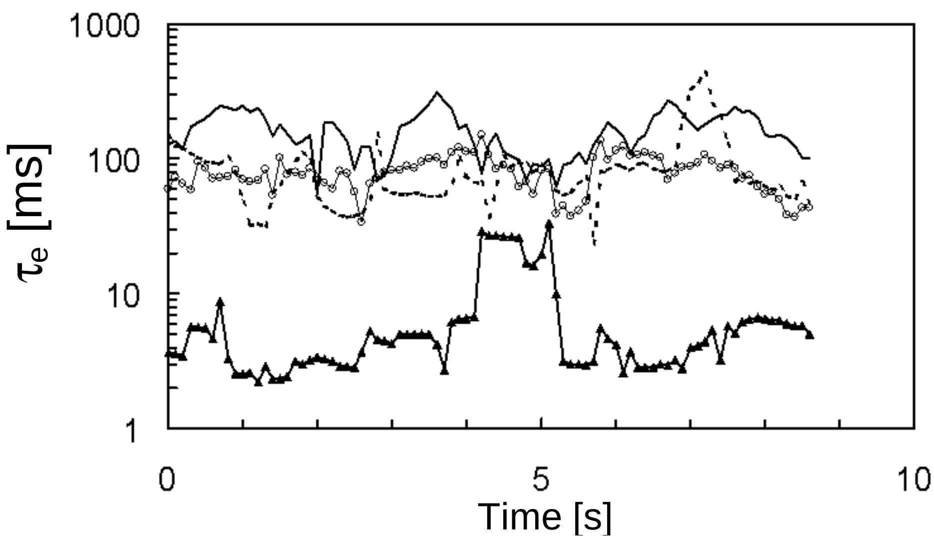

Fig. 2 Examples of τ_e extracted from the running ACF of motif I (FP) sung by subject KK using four different lyrics. ( ): humming, $(\tau_e)_{min}$ = 54 ms; (○): sung with "la" syllables, $(\tau_e)_{min}$ = 34 ms; (----): sung with real lyrics, $(\tau_e)_{min}$ = 23 ms; (▲): non-voiced singing, $(\tau_e)_{min}$ = 2.2 ms.

Table 1 Measured average values of $(\tau_e)_{min}$ and its standard deviation for a music motif sung by each subject using four different lyrics.

	$(t_e)_{min}$ [ms]				
	CT (Sop.)	YA (Alt.)	KK (Ten.)	DT (Bas.)	**Global**
humming	73 ± 14	51 ± 13	64 ± 18	70 ± 20	**65 ± 18**
la	49 ± 6	36 ± 5	34 ± 11	26 ± 3	**36 ± 11**
lyrics	33 ± 8	28 ± 6	14 ± 4	18 ± 2	**23 ± 10**
non-voiced		2.1 ± 0.6	2.0 ± 0.2		**2.0 ± 0.4**

III. EXPERIMENT WITH DIFFERENT MUSIC TEMPOS

PROCEDURES

Although there was preliminary indication that tempo was less important for temporal variation than the variation of speech in lyrics [5], we assumed that values of $(\tau_e)_{min}$ for vocal performances would differ with music tempo as was observed with piano performances [6]. In order to systematically examine the difference of the $(\tau_e)_{min}$ when the motif were sung in different music tempo, an analysis of $(\tau_e)_{min}$ of the sound source with four singers for two music motifs was conducted.

Subjects who participated in the study with different lyrics were asked to sing motif I (FP) and motif I (CP) with "la" syllables. For the ease of playing, the range of notes was adjusted for each singer as in the previous experiment. The music tempo were ♩ = 60, 120, and 180 for the two motifs.

RESULTS

The mean values and standard deviations of $(\tau_e)_{min}$ obtained in ten trials for the two motifs played with the three different music tempos are listed in Table 2.

Values of $(\tau_e)_{min}$ did not differ with music tempo ($p < 0.05$) except for one case [subject YA, motif I (FP), ♩ = 60]. Values of $(\tau_e)_{min}$ of motif I (CP) were longer than those of motif I (FP) ($p < 0.05$) except for one case [subject CT, ♩ = 180].

Table 2 Measured average values of $(\tau_e)_{min}$ and its standard deviation for two music motifs played by each subject with three different music tempos.

		$(t_e)_{min}$ [ms]				
Motif	♩	CT	YA	KK	DT	**Global**
	60	54 ± 11	25 ± 4	42 ± 6	27 ± 5	**36 ± 15**
Motif I	120	49 ± 6	36 ± 4	34 ± 11	26 ± 3	**36 ± 11**
(FP)	180	59 ± 9	32 ± 9	39 ± 13	31 ± 9	**40 ± 15**
	60	65 ± 11	43 ± 7	51 ± 5	35 ± 3	**49 ± 14**
Motif I	120	58 ± 9	48 ± 12	49 ± 9	40 ± 9	**49 ± 11**
(CP)	180	58 ± 13	51 ± 14	56 ± 12	42 ± 5	**52 ± 13**

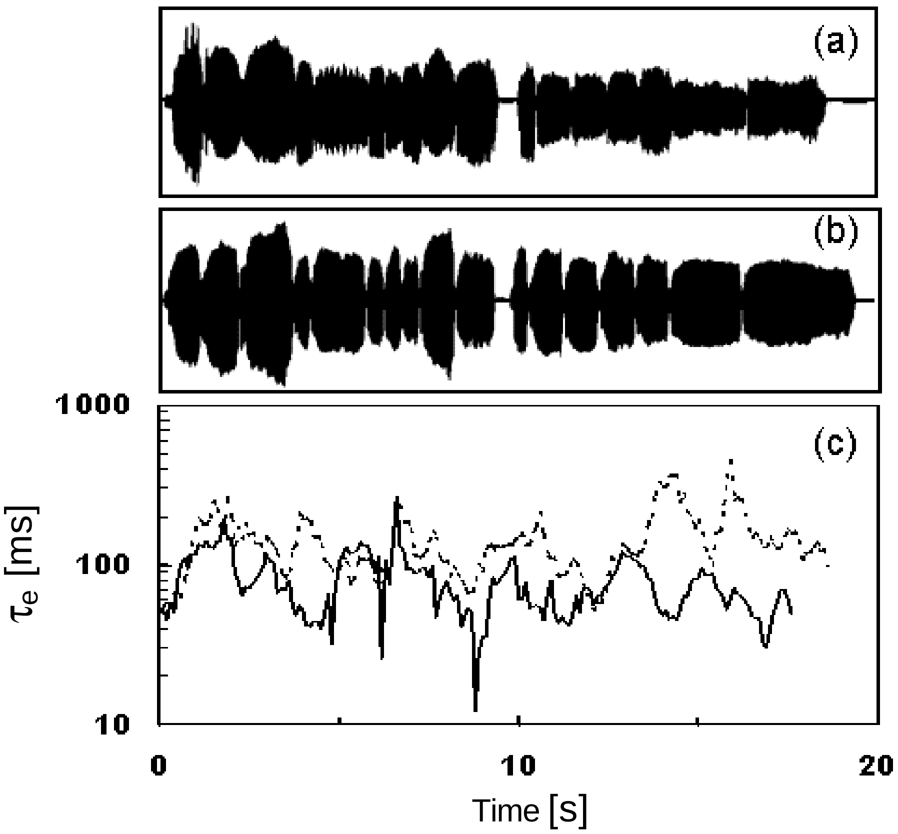

Fig. 3 Examples of real waveform and their value of τ_e for Motif I (FP) with ♩ = 60 played by subject YA and subject KK. (a): Waveform for subject YA. (b) Waveform for subject KK. (c) Example of τ_e for the two subjects. (----): Subject KK, $(\tau_e)_{min}$= 55 [ms]; ( ): Subject YA, $(\tau_e)_{min}$= 30 [ms].

IV. DISCUSSION

The results of Table 1 showed that the values of $(\tau_e)_{min}$ for each subjects are relatively longer in order of: humming > singing "la" > singing with lyrics > non-voiced singing ($p < 0.01$). Like humming singing, the value of τ_e of any source signal containing low levels of noise components will be long.

The results of Table 2 showed that the values of $(\tau_e)_{min}$ in vocal performances are not dependent on the music tempo ($p < 0.05$). This is unlike the case of piano performances [6] and may be due to the following cancellation effect: in singing fast tempo music, the fluctuation of the pitch becomes larger while the fluctuation of the energy envelope becomes smaller.

The results of Table 2 also showed that the fluctuation of pitch is an important element to determine the value of $(\tau_e)_{min}$ ($p < 0.05$).

The value of $(\tau_e)_{min}$ of motif I (FP) played by subject YA with ♩ = 60 was much lower than the other cases, which may be due to the effect of vibrato. Fig. 3 (a) shows a typical waveform of the source sung by the subject YA. The fluctuations of envelope due to vibrato can be seen around the period of both 5 s and 18 s, comparing to the waveform of the source sung by subject KK shown in Fig. 3 (b). Since the low τ_e values between the periods of 8-9 s are considered to be affected by breath noise, they were passed over in the analysis to determine the minimum value.

REFERENCES

1. Y. Ando 1998 Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners. AIP/Springer-Verlag, New York.

2. Y. Ando, T. Okano, Y. Takezoe 1989 J. Acoust. Soc. Am., **86**, 644-649. The running autocorrelation function of different music signals relating to preferred temporal parameters of sound field.

3. I. Nakayama 1984 Acustica **54**, 217-221. Preferred time delay of a single reflection for performers.

4. S. Sato, Y. Ando, and S. Ota 2000 Journal of Sound and Vibration **232**, 27-37. Subjective preference of cellists for the delay time of a single reflection in a performance.

5. D. Noson, S. Sato, H. Sakai and Y. Ando 2000 Journal of Sound and Vibration **232**, 39-51. Singer responses to sound fields with a simulated reflection.

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

7. K. Tom and I. Titze 2001 J. Acoust. Soc. Am., **110**, 1667-1676. Vocal intensity in falsetto phonation of a countertenor: An analysis by synthesis approach.

8. J. Sundberg 1987 Northern Illinois University Press, DeKalb, IL. The science of the singing voice.

9. S. Ternstöm and J. Sundberg 1989 J. Acoust. Soc. Am., **86**, 517-522. Formant frequencies of choir singers.

10. A. Gade 1989 Acustica **54**, 217-221. Investigations of musician's room acoustics conditions in concert halls, Part II.

11. A. H. Marshall and J. Meyer 1985 Acustica **58**, 130-140. The directivity and auditory impression of singers.

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

13. K. Kato and Y. Ando 2001 17th ICA Proceedings Volume III. Blending voice source and temporal factor of sound field by selecting different singing types, submitted to Journal of Sound and Vibration (*to be published in 2002*).

14. K. Mouri and K. Akiyama 2001 Journal of Sound and Vibration **241**, 87-95. Preliminary study on recommended time duration of source signals to be analyzed, in relation to its effective duration of the auto-correlation function.

APPENDIX. $(\tau_e)_{min}$: A TYPICAL TEMPORAL PARAMETER OF SOURCE SIGNAL

The effective duration of ACF is defined by the delay τ_e at which the envelope of the ACF becomes –10 dB (see Fig. A1). The minimum value of the fluctuation of τ_e, $(\tau_e)_{min}$, is the most active part of each piece, influencing subjective preference for both listeners and performers.

For listeners, the most preferred initial time delay $[\Delta t_1]_p$ and the most preferred subsequent reverberation time $[T_{sub}]_p$ are expressed by

$$[\Delta t_1]_p \approx (1-\log_{10}A)(\tau_e)_{min} \quad (1)$$

$$[T_{sub}]_p \approx 23\,(\tau_e)_{min} \quad (2)$$

where A is the total amplitude of the reflection [1, 2].

For performers, the most preferred delay time of a single reflection t_p is described by

$$t_p = \left(\log_{10}\frac{1}{k} - c\log_{10}A\right)(t_e)_{min} \quad (3)$$

where the values k and c are constants that depend on the subjective attributes [1, 3, 4].

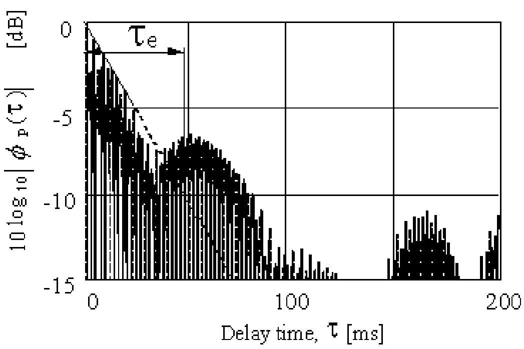

Fig. A1 Example of determining the effective duration of the running ACF. τ_e = 48 ms.