*Journal of Sound and Vibration* (2002) **258**(3), 463–472
doi:10.1006/jsvi.5269, available online at http://www.idealibrary.com on IDEAL®

# A STUDY OF THE BLENDING OF VOCAL MUSIC WITH THE SOUND FIELD BY DIFFERENT SINGING STYLES

K. Kato and Y. Ando

*Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe* 657-8501, *Japan. E-mail: kkato@kobe-u.com*

(*Accepted* 30 *May* 2002)

The initial time delay $\Delta t_1$ 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 do not have ideal conditions for all forms of music. It has been shown that the most preferred conditions for both listeners and performers are determined by the minimum value of the effective duration of the running autocorrelation function (ACF) of sound signals, $(\tau_e)_{min}$. To determine the suitability of vocal music for a given sound field, $(\tau_e)_{min}$ of vocal music was analyzed, after recording five solo singers (tenor) in an anechoic room. The results showed that $(\tau_e)_{min}$ of the ACF of a voice source, which is closely related to the two temporal factors of the sound field, varies with singing style. A significant finding is that the values of $(\tau_e)_{min}$ of sound signals for *falsetto* and *medium falsetto* are significantly longer than that for *operatic singing*.



## 1. INTRODUCTION

Opera houses and concert halls have unique characteristics. Obviously, no hall is suitable for playing all kinds of music. Thus, it would be useful to find measures of this diversity. The decay characteristic of the autocorrelation function (ACF) of a sound source is closely related to the subjective preference of sound fields. It was reported that the minimum value of the effective duration of running ACF, $(\tau_e)_{min}$, is a typical temporal parameter of a sound source and that it is closely related to the most-preferred temporal conditions for both listeners [1–4] and performers [5–7]. In the context of this theory, it is important to estimate $\tau_e$ of running ACF of music. Taguti and Ando [8] revealed that tempo, articulation, and damper pedalling are the main elements for determining $\tau_e$ of piano music. In contrast, the present study deals with vocal performances and examines the effect of different interpretation styles on the behavior of running ACF. It is worth noting that the factors extracted from the power spectrum may not be good for describing a temporal factor for blending source signals and a given sound field (see Appendix A).

## 2. PROCEDURE

### 2.1. DIFFERENT SINGING STYLES

Factors of vocal music assumed to characterize $\tau_e$ of running ACF are: singing style, lyrics, note value or tempo, dynamics such as crescendo and decrescendo, various kinds of articulation such as staccato, legato (or tenuto), marcato, sforzando, etc.; and others such

 


Figure 1. Music scores of the four motifs used in the experiments.

as vibrato or tremolo. The singing styles of *falsetto*, *medium falsetto*, and *operatic singing* were studied. *Falsetto* is considered to be the most typical singing styles for increasing $\tau_e$, while *operatic singing* is considered to decrease $\tau_e$. *Medium falsetto* is defined as a mode between *falsetto* and *operatic singing*.

## 2.2. MUSIC SCORES

Figure 1 shows the four musical scores used in this experiment. E4 ‘‘oh’’ vowel and motif I are examples of motifs with a simple structure. Motif I consists of eighth note and arranged longer in horizontal intervals of major and minor seconds, with no indication of expression marks such as staccato, stress accent, etc. Therefore, it is expected that the $\tau_e$ values of the voice source of these two motifs mainly depend on the singing style. On the other hand, motifs II and III contain rather complex structures. Motif II has sixteenth note. Motif III has staccatos and accents as well as sixteenth note and thirty-second note. Thus, the $\tau_e$ values of the voice source can be characterized by the singing style and these factors.

E4 ‘‘oh’’ vowel was sung by each subject 10 times in *falsetto*, and then sung 10 times in *operatic singing* style. For music pieces I–III, each motif was sung 10 times in *medium falsetto*, and then sung 10 times in *operatic singing* style.

### 2.3. CONDITIONS

The singer's voice was picked up by a microphone located $25 \pm 1$ cm in front of the singer and $5 \pm 0{\cdot}5$ cm to the side of the singer's mouth in an anechoic room. Signals were sampled at a 44·1 kHz. The music tempo was maintained by the aid of a visual metronome located in front of the singer.

## 3. RESULTS

### 3.1. PRIMARY STUDY WITH ONE SINGER

A primary analysis of $(\tau_e)_{min}$ of the sound source with a solo singer (tenor) for four music motifs was conducted to examine the difference of the $(\tau_e)_{min}$ when the four motifs were sung in different singing styles.

Figure 2 shows the measured $\tau_e$ values of the running ACF with a 100-ms interval as a function of time for ten trials. As the recommended integration interval $(2\,T)_r$ is around $30(\tau_e)_{min}$ [9], $2\,T = 2{\cdot}0$ s was selected for E4 "oh" vowel, while $2\,T = 1{\cdot}0$ s was selected for motifs I–III. Figures 2(a) and 2(b) show that $\tau_e$ for E4 "oh" vowel was constant in each session. On the other hand, $\tau_e$ for real motifs I–III fluctuated in each session. Figures 2(c) and 2(d) indicate that $\tau_e$ for motif I was shorter during the 1–1·5 s part than in the other parts. Figures 2(e) and 2(f) show that $\tau_e$ for motif II was shorter during the 0–0·5 s or 1·8–2·2 s parts than in the other parts. Figures 2(g) and 2(h) show that fluctuations of $\tau_e$ were different in each session.

Figure 3 shows the measured $(\tau_e)_{min}$ of the running ACF obtained in 10 trials for the four motifs. The mean values and standard deviations of $(\tau_e)_{min}$ are listed in Table 1. Analyzing the variance showed that $(\tau_e)_{min}$ for *falsetto* or *medium falsetto* is significantly ($p < 0{\cdot}01$) longer than that for *operatic singing* of the E4 "oh" vowel, motif I, and motif III, but not for motif II. The $(\tau_e)_{min}$ varied with singing style, especially for the simple motifs (E4 "oh" vowel and motif I). It was expected that a singer controls $\tau_e$ for simple motifs with the chosen singing, even if his chorus experience was 10 years.

### 3.2. EXPERIMENTS WITH E4 "OH" VOWEL AND MOTIF I

Experiments using five singers (tenor) were conducted on $(\tau_e)_{min}$ of the sound source in order to examine individual differences. Following the above results, only E4 "oh" vowel and motif I were chosen in this experiment.

Figure 4 shows the range of $(\tau_e)_{min}$ for the two motifs for each trial. One-dimensional variance analysis showed significant differences ($p < 0{\cdot}05$) of $(\tau_e)_{min}$ between the two singing styles, except for one case (Subject YN, motif I). For E4 "oh" vowel, the average $(\tau_e)_{min}$ for *falsetto* was 258 ms, while the average for *operatic singing* was 95 ms. For motif I, the average of the $(\tau_e)_{min}$ for *medium falsetto* was 39 ms, while the average for *operatic singing* was 29 ms.

Four of the five singers could control $(\tau_e)_{min}$ of the running ACF of E4 "oh" vowel and motif I by varying the singing styles.

## 4. DISCUSSION

Four of the five singers whose chorus experiences were within 10 years could control $(\tau_e)_{min}$ for simple motifs in *falsetto* and *operatic singing* style of E4 "oh" vowel, and

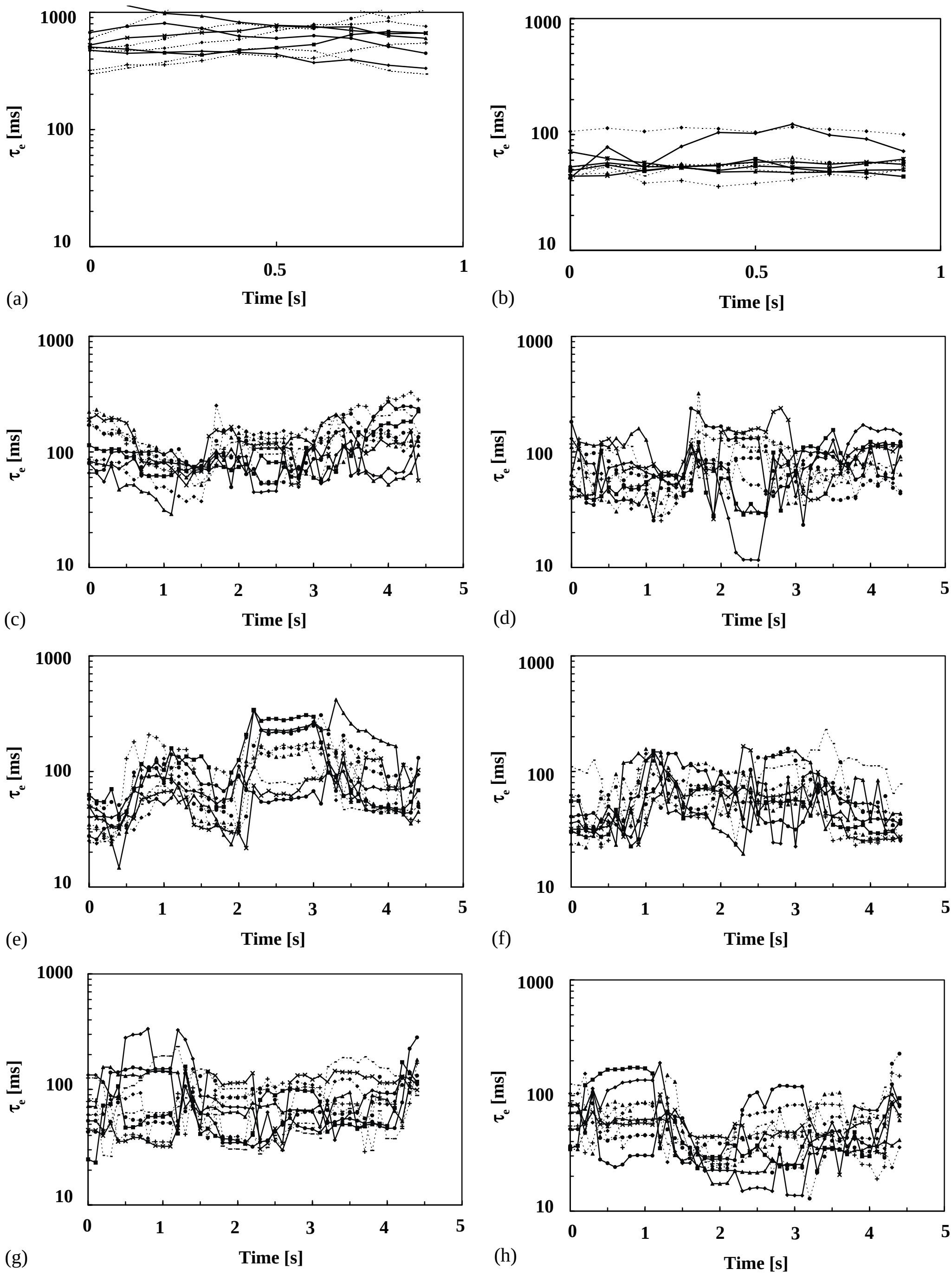


Figure 2. Examples of $\tau_e$ extracted from the running ACF as a function of time for 10 trials of E4 "oh" vowel and motifs I–III with a single subject. (—◆—): first session; (—■—): second session; (—▲—): third session; (—×—): fourth session; (—●—): fifth session; (- - -● - - -): sixth session; (- - -+ - - -): seventh session; (- - —- -): eighth session; (- - - ▲- - -): ninth session; (- - - ◆- - - ): tenth session. For E4 "oh" vowel, integration time: $2T = 2{\cdot}0$ s, running step: 100 ms, total signal duration of each sample: 3·4 s. For motifs I–III, integration time: $2T = 1{\cdot}0$ s, running step: 100 ms, and total signal duration of each sample: 5·4 s. (a) E4 "oh" vowel, *falsetto*, $(\tau_e)_{min} = 452 \pm 109$ ms. (b) E4 "oh" vowel, *operatic singing*, $(\tau_e)_{min} = 47 \pm 10$ ms. (c) Motif I, *medium falsetto*, $(\tau_e)_{min} = 49 \pm 10$ ms. (d) Motif I, *operatic singing*, $(\tau_e)_{min} = 27 \pm 7$ ms. (e) Motif II, *medium falsetto*, $(\tau_e)_{min} = 27 \pm 8$ ms. (f) Motif II, *operatic singing*, $(\tau_e)_{min} = 25 \pm 6$ ms. (g) Motif III, *medium falsetto*, $(\tau_e)_{min} = 32 \pm 7$ ms. (h) Motif III, *operatic singing*, $(\tau_e)_{min} = 20 \pm 5$ ms.

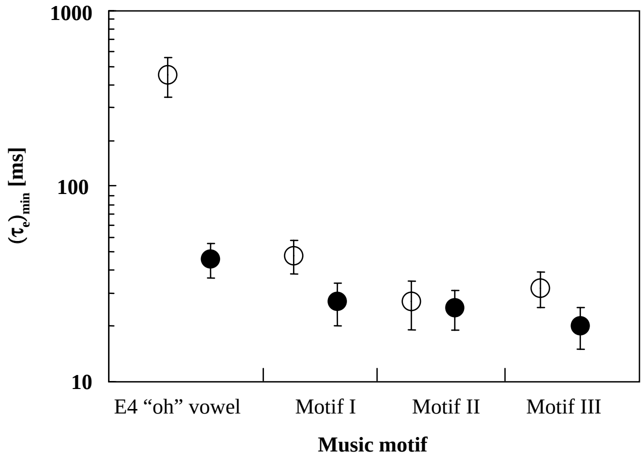


Figure 3. Examples of averaged values of $(\tau_e)_{min}$ and their standard deviations for 10 trials (subject AH). ○, *falsetto* (E4 "oh" vowel) or *medium falsetto* (motifs I–III). ●, *operatic singing.*

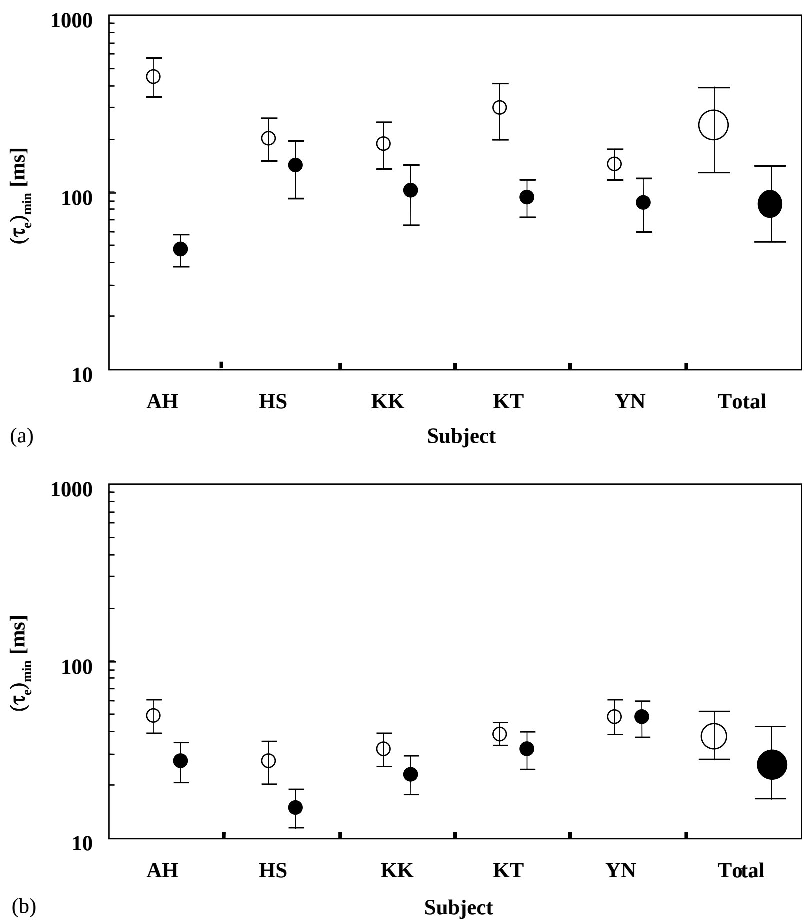


Figure 4. Averaged values of $(\tau_e)_{min}$ and their standard deviations for 10 trials for each subject. (a) E4 "oh" vowel (○, *falsetto*, ●, *operatic singing*), and (b) Motif I (○, *medium falsetto*; ●, *operatic singing*).

TABLE 1

*Example of measuring averaged values of $(\tau_e)_{min}$ and its standard deviation for four music pieces played by a single subject (AH) using two different singing styles. Great differences due to singing style for both E4 "oh" vowel and motif I with a simple structure are evident*

| Source | *Falsetto or medium falsetto* | *Operatic singing* (ms) | *P* |
|---|---|---|---|
| E4 "oh" vowel | $452 \pm 109$ | $47 \pm 10$ | $<0{\cdot}01$ |
| Motif I | $49 \pm 10$ | $27 \pm 7$ | $<0{\cdot}01$ |
| Motif II | $27 \pm 8$ | $25 \pm 6$ | — |
| Motif III | $32 \pm 7$ | $20 \pm 5$ | $<0{\cdot}01$ |

*medium falsetto* and *operatic singing* style of motif I. Therefore, for simple motifs, the singing style affects the temporal factor of the sound source, $(\tau_e)_{min}$. *Falsetto* or *medium falsetto* increases $(\tau_e)_{min}$, and *operatic singing* decreases $(\tau_e)_{min}$.

The results of Figure 3 showed a lower degree of reproducibility of $\tau_e$ than that in a cello performance [6]. This might be due to the singers' lack of skill. A singer who trained more may be better able to reproduce instances of song than the present subjects. Figure 4 indicates that the degree of fluctuation in $(\tau_e)_{min}$ of the two singing styles may depend on the subject's skill. Figures 2(c) and 2(d) indicate that double consonants such as "sf" may decrease $\tau_e$. Figures 2(e) and 2(f) indicate that consonants such as "p" and "v" may decrease $\tau_e$. Figures 2(g) and 2(h) indicate that motifs with staccato and accent may be difficult to reproduce.

## 5. CONCLUSIONS

It was found that the style of singing affects the typical temporal factor of the sound signal, i.e., $(\tau_e)_{min}$. Values of $(\tau_e)_{min}$ of voice sources for *falsetto* or *medium falsetto* are much longer than that for *operatic singing*.

To estimate $\tau_e$ of singing voices, further experiments are needed to examine other factors: lyrics, note value or tempo, 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.

## ACKNOWLEDGMENTS

The authors thank Akira Hirabayashi for his useful suggestions about the choice of musical score and the singers who participated in the experiments.

## REFERENCES

1. Y. ANDO 1998 *Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners*. New York: AIP/Springer-Verlag.
2. Y. ANDO 1977 *Journal of the Acoustical Society of America* **62**, 1436–1441. Subjective preference in relation to objective parameters of music sound fields with a single echo.
3. Y. ANDO and K. KAGEYAMA 1977 *Acustica* **37**, 111–117. Subjective preference of sound with a single early reflection.

4. Y. ANDO, T. OKANO and Y. TAKEZOE 1989 *Journal Acoustical Society of America* **86**, 644–649. The running autocorrelation function of different music signals relating to preferred temporal parameters of sound field.
5. I. NAKAYAMA 1984 *Acustica* **54**, 217–221. Preferred time delay of a single reflection for performers.
6. 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.
7. 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.
8. T. TAGUTI and Y. ANDO 1997 in *Music and Concert Hall Acoustics.* (Y. Ando and D. Noson, editors) Academic Press; London: Chapter 23. Characteristics of the short-term autocorrelation function of sound signals in piano performances.
9. 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.
10. S. TERNSTÖM and J. SUNDBERG 1989 *Journal of the Acoustical Society America* **86**, 517–522. Formant frequencies of choir singers.
11. K. TOM and I. TITZE 2001 *Journal of the Acoustical Society America* **110**, 1667–1676. Vocal intensity in falsetto phonation of a countertenor: an analysis by synthesis approach.
12. G. BLOOTHOOFT, E. BRINGMANN, M. CAPPELLEN, J. LUIPEN and K. THOMASSEN 1992 *Journal of the Acoustical Society America* **92**, 1827–1836. Acoustics and perception of overtone singing.
13. H. SMITH, K. STEVENS and R. TOMLINSON 1967 *Journal of the Acoustical Society America* **41**, 1262–1264. On an unusual mode of chanting by certain Tibetan lamas.
14. J. SUNDBERG 1987 *The Science of the Singing Voice*. DeKalb, IL: Northern Illinois University Press.

## APPENDIX A: EXAMPLES OF REAL WAVEFORM, POWER SPECTRUM AND ACF

Figure A1 shows examples of the real waveform, the power spectrum and the ACF of the voice source of the three singing styles. Numerous studies have been reported on the power spectrum of voice sources [10–14]. For example, the power spectrum reveals that the voice source in *falsetto* or *medium falsetto*, like *choral singing* analyzed by Ternström and Sundberg [10], apparently contains less singer's formant (2–3·5-kHz spectrum envelop) and fewer high frequency components than does *operatic singing*. However, it is hard to extract useful temporal parameters from the running power spectrum. $\tau_e$ directly correlates with preferred temporal parameters of the sound field such as $[\Delta t_1]_p$ or $[T_{sub}]_p$. Thus, this temporal factor containing source signals can be used to blend the sound source with the sound field.

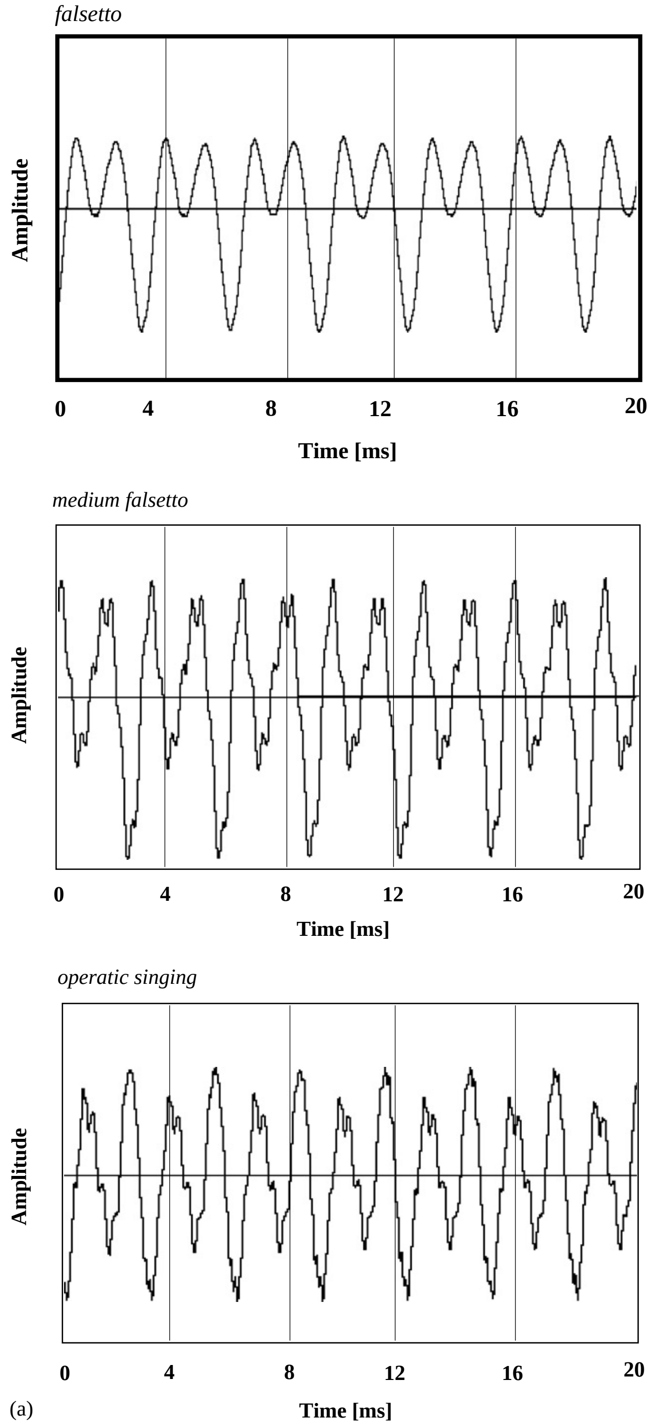


Figure A1. Examples of (a) real waveform, (b) power spectrum, and (c) ACF of three singing styles: *falsetto*, *medium falsetto*, and *operatic singing* (singing E4 "oh" vowel).

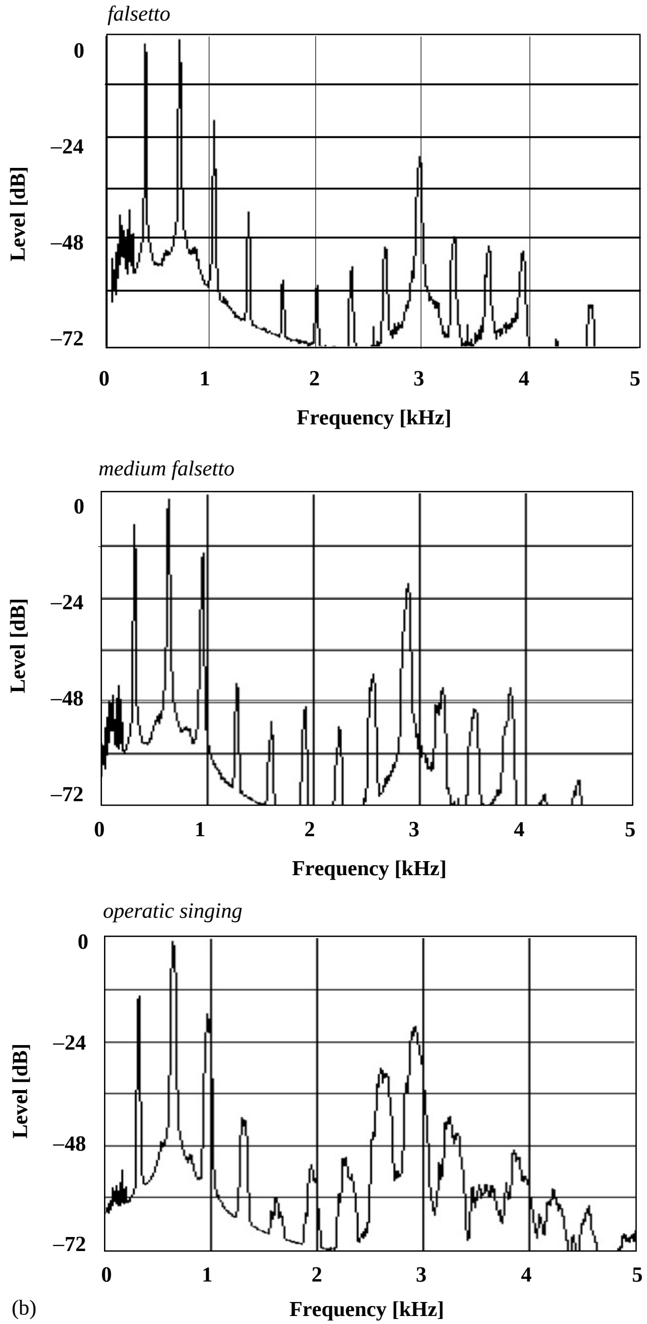


Figure A1. Continued.

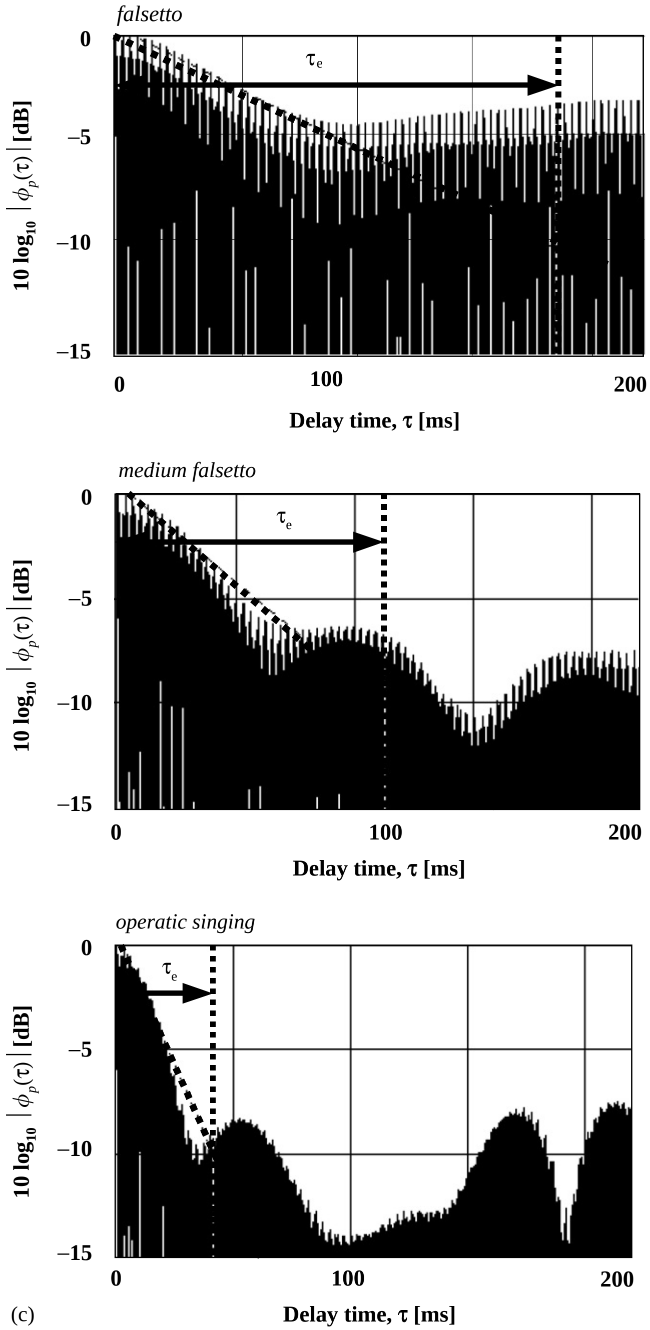


Figure A1. Continued.