Journal of Sound and Vibration (2001) **241**(1), 87–95
doi:10.1006/jsvi.2000.3280, available online at http://www.idealibrary.com on IDEAL®

PRELIMINARY STUDY ON RECOMMENDED TIME DURATION OF SOURCE SIGNALS TO BE ANALYZED, IN RELATION TO ITS EFFECTIVE DURATION OF THE AUTO-CORRELATION FUNCTION

K. MOURI AND K. AKIYAMA

Central Research & Development Headquaters, Itoki Crebio Corp., 4-12, *Imafuku-higashi*, 1-*chome*, *Joto-ku*, *Osaka* 536-0002, *Japan*

AND

Y. ANDO

Graduate School of Science and Technology, *Kobe University*, *Rokkodai*, *Nada*, *Kobe* 657-8501, *Japan*

(*Accepted* 7 *August* 2000)

A workable model of human auditory–brain system including the auto-correlation function (ACF) and the interaural cross-correlation function (IACF) has been reported [1]. Subjective evaluations can be described by physical factors extracted from both the ACF of sound source and the IACF. The purpose of this study is to find a recommended integration time $(2T)_r$ by using several noise and sound sources existing in an office. Results show the most suitable integration interval from the viewpoints of the loudness and the pitch can be determined in relation to the effective duration (τ_e) of the sound source.

1. INTRODUCTION

Subjective evaluations for the environmental sound should be considered to design the sound environment for each office worker. In the concert hall acoustics, the human auditory–brain system characterized by the auto-correlation function (ACF) and the interaural cross-correlation function (IACF) and the specialization of cerebral hemispheres has been well accepted for the subjective evaluations [1–4]. In order to calculate the running ACF of music and speech signal for subjective preference judgement, the integration time duration $(2T)$ has been selected as 2·0 s. In this paper, we discuss a suitable $(2T)$ for noise and other sounds evaluating the loudness and the pitch. Such a duration of $(2T)$ is considered as a time window of our auditory–brain system because it is regarded as minimum unit for analyzing sound sources.

The purpose of this study is to research the $2T$ relating to the loudness and the pitch of noise sources in an office.

2. DEFINITION OF THE PHYSICAL FACTORS OF THE ACF AND THE IACF

The physical factors, $\Phi(0)$, ϕ_1, τ_e, and τ_1 were extracted from the ACF and IACC, τ_{IACC}, and W_{IACC} were extracted from the IACF [1, 3, 4]. The definition of the ACF and the IACF

0022-460X/01/110087 + 09 $35.00/0

are expressed as follows:

$$\Phi_p(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{+T} p'(t)p'(t+\tau)\,\mathrm{d}t, \quad \Phi_{lr}(t) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{+T} p'_l(t)p'_r(t+\tau)\,\mathrm{d}t.$$

Here $p'(t) = p(t)*s(t)$, $s(t)$ being the ear sensitivity. For convenience, $s(t)$ may be chosen as the impulse response of an A-weighted network.

$\Phi_{lr}^{(n)}(\tau)$ is the IACF of the nth reflection; $\Phi_{ll}^{(n)}(0)$, and $\Phi_{rr}^{(n)}(0)$ are the respective sound energies arriving at the two ears from the nth reflection.

Four physical factors are obtained from the ACF: (1) energy represented at the origin of the delay, $\Phi_p(0)$; (2) effective duration of the envelope of the normalized autocorrelation function, τ_e (which is defined by the 10-percentile delay), representing a kind of repetitive feature or reverberation containing the source signal itself; (3, 4) fine structure, including peaks and dips with their delays: the delay time and amplitude of the first peak–namely, τ_1 and ϕ_1.

Three physical factors are obtained from the IACF: (1) IACC is the maximum value of the IACF within the time delay of 1 ms; (2) τ_{IACC} corresponds to the interaural time delay for the horizontal angle; (3) W_{IACC} corresponds to the apparent source width (ASW) mainly.

In this research, $\Phi(0)$ is associated with the loudness and the pitch is considered to be determined by ϕ_1 and τ_1.

3. SOUND ANALYSIS

In this research, telephone ringing, the fan noise of the air conditioner, the sound of a key-punch, the fan noise of the personal computer, voice, music H (Beethoven, Symphonie No. 6, F - dur, op. 68, Pastorale Andante molto mosso), and music I (Mozart, String Quintet, No. 4, first movement) were chosen as the sound sources. Music H has a slow tempo while music I has a fast tempo. These music sounds were regarded as background music (BGM).

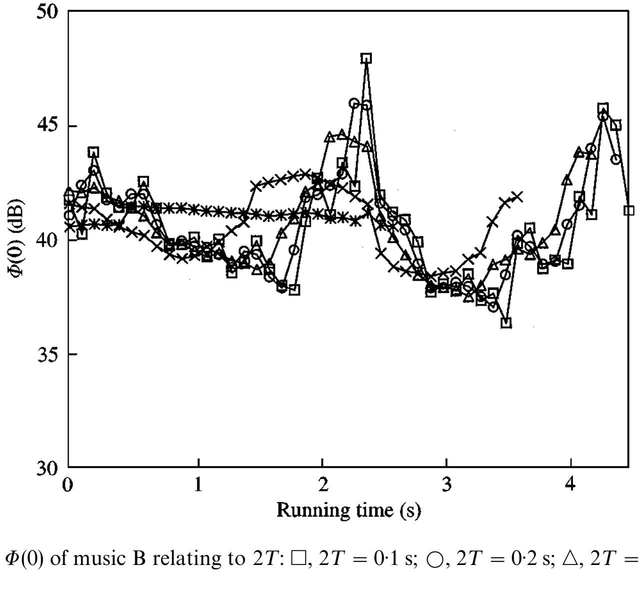

Figure 1. Running $\Phi(0)$ of music B relating to $2T$: □, $2T = 0{\cdot}1$ s; ○, $2T = 0{\cdot}2$ s; △, $2T = 0{\cdot}4$ s; ×, $2T = 1{\cdot}0$ s; ✱, $2T = 2{\cdot}0$ s.

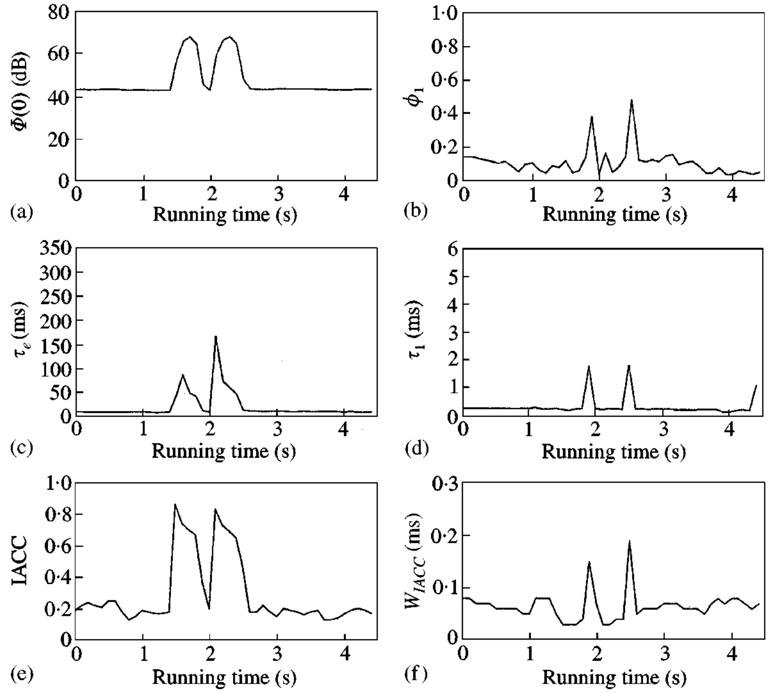

Figure 2. Analysis of the running ACF and the running IACF of telephone ringing with $2T = 0{\cdot}2$ s: (a), $\Phi(0)$; (b), ϕ_1; (c), τ_e; (d), τ_1; (e), IACC; (f), W_{IACC}.

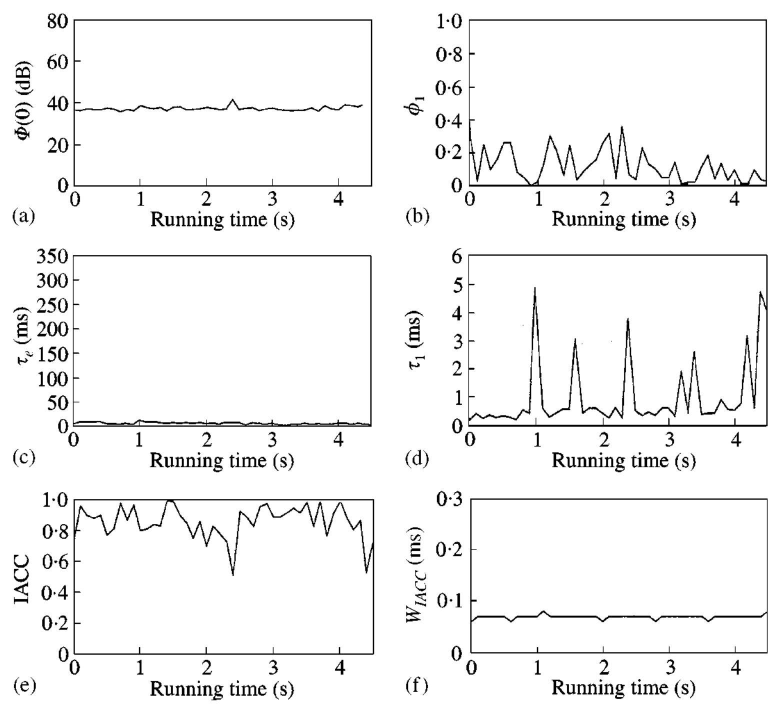

Figure 3. Analysis of the running ACF and the running IACF of the noise of air conditioner with $2T = 0{\cdot}1$ s. (a)–(f) as Figure 2.

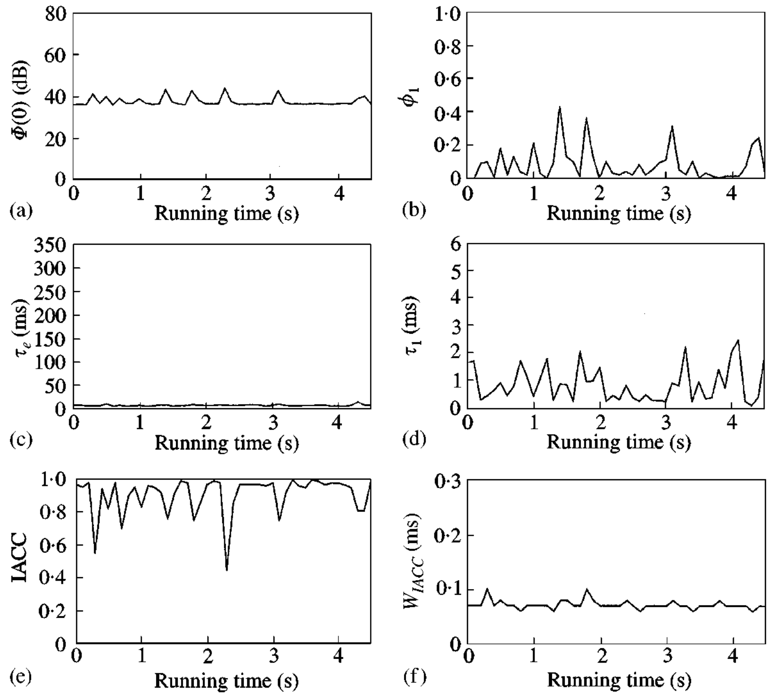

Figure 4. Analysis of the running ACF and the running IACF of the noise of key-punch with $2T = 0{\cdot}1$ s. (a)–(f) as Figure 2.

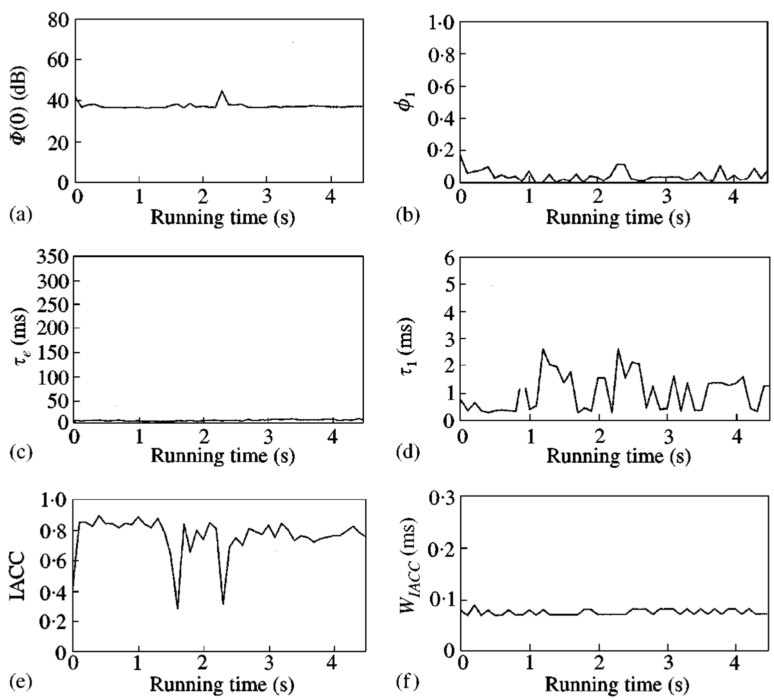

Figure 5. Analysis of the running ACF and the running IACF of fan noise of personal computer with $2T = 0{\cdot}1$ s. (a)–(f) as Figure 2.

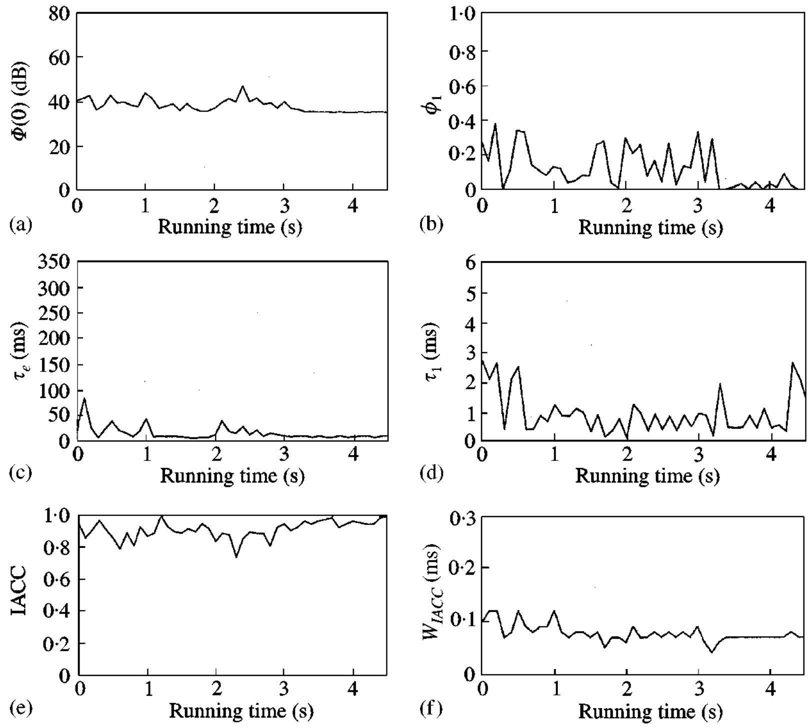

Figure 6. Analysis of the running ACF and the running IACF of voice with $2T = 0{\cdot}1$ s. (a)–(f) as Figure 2.

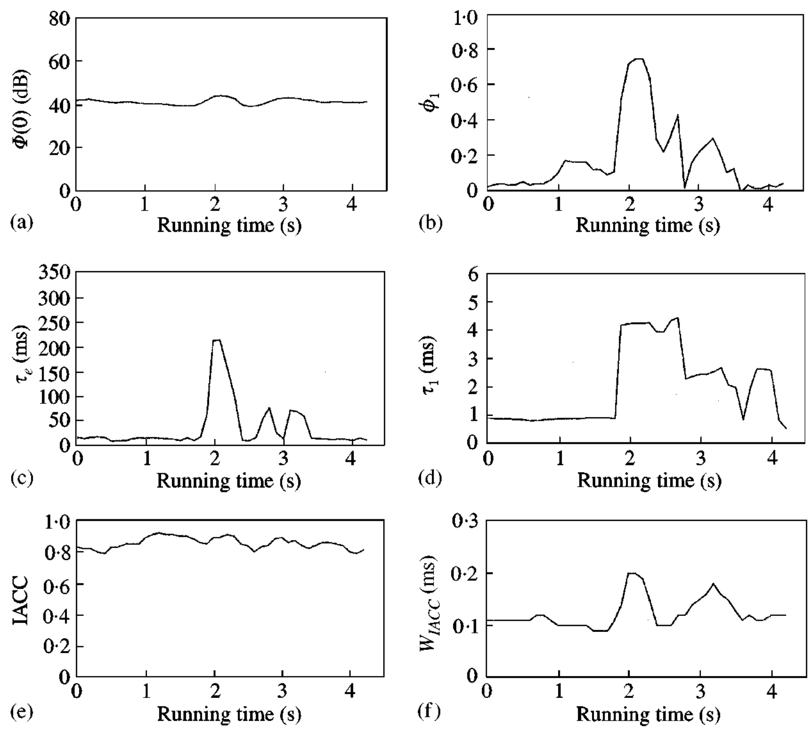

Figure 7. Analysis of the running ACF and the running IACF of music H with $2T = 0{\cdot}4$ s. (a)–(f) as Figure 2.

The running evaluated physical factors of the ACF and the IACF of telephone ringing was calculated for $2T = 0{\cdot}1$, 0·2, 0·4, 1·0 and 2·0 s with the running interval of 0·1 s, as shown in Figures 1–7 because the frequency information up to 10 Hz is included when $2T$ is up to 0·1 s.

Firstly, the correlation coefficients among each physical factors ($\Phi(0)$, ϕ_1, τ_e, and τ_1) were investigated. Secondly, the waveforms of $\Phi(0)$, ϕ_1, τ_e, and τ_1 for each $2T$ was compared to its hearing impression, respectively, and the $2T$'s, which corresponded to the hearing impression of the waveform, were selected for each sound sources. Thirdly, the $2T$'s were investigated according to the minimum value of effective duration $(\tau_e)_{min}$ because $(\tau_e)_{min}$ is considered to be one of the most important factors for perception of timbre [2]. This $2T$ is expressed as $(2T)_r$.

These analyses were carried out by use of the Real Time Sound Analyzer (Yoshimasa Electronic Inc.). Data were recorded from two microphones fixed at the left and right ears of a real head with the sampling rate of 44·1 kHz.

4. RESULTS AND DISCUSSION

First of all, the correlation coefficients of $\Phi(0)$, ϕ_1, τ_e, and τ_1 of voice and fan noise of the air conditioner and personal computer are low, because the factors are theoretically orthogonal with each other. But, those of the key-punch, telephone ringing, music H and music I are apparently high; for example, the value for music I between $\Phi(0)$ and ϕ_1 is 0·82 because of its rhythmical sound, particularly with respect to τ_e and τ_1 (see Table 1). The relationship between waveform and hearing impression was observed in $\Phi(0)$ and τ_1, but was not found in ϕ_1 and τ_e, because these factors might be masked by $\Phi(0)$ and τ_1. A typical example is shown in Figure 1 and Table 2. Additionally, the recommended $2T$'s of $\Phi(0)$ and

TABLE 1

Correlation coefficients above 0·7 among physical factors

Sound source	Physical factors	Correlation coefficient
Key-punch	$\Phi(0)$ and ϕ_1	0·75
Music H	ϕ_1 and τ_1	0·79
	ϕ_1 and τ_e	0·86
Music I	$\Phi(0)$ and ϕ_1	0·82
	$\Phi(0)$ and τ_e	0·82
Telephone ringing	$\Phi(0)$ and τ_e	0·77
	ϕ_1 and τ_1	0·78

TABLE 2

Example of judgement of hearing impression for several $2T$s (Music B)

$2T$ (s)	0·1	0·2	0·4	1·0	2·0
Viewpoint $\Phi(0)$	△	○	△	×	×

Note: (○), well matched; (△), mildly matched; (×), little matched.

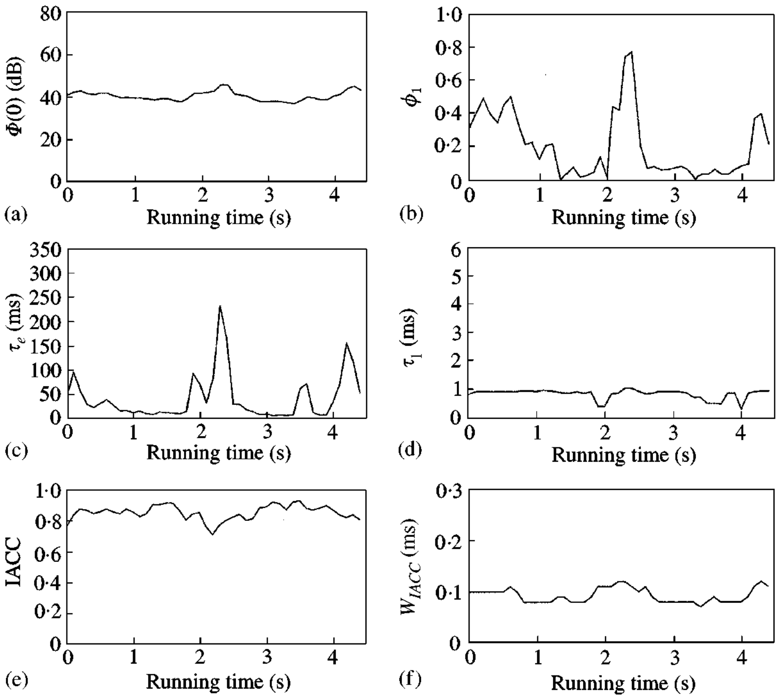

Figure 8. Analysis of the running ACF and the running IACF of music I with $2T = 0{\cdot}2$ s. (a)–(f) as Figure 2.

τ_1 corresponded to each other. For the fan noise of the air conditioner, the key-punch noise, the fan noise of the personal computer and the human voice, the value of $(2T)_r$ is recommended as 0·1 s. For the telephone ringing and music I $2T \approx 0{\cdot}2$ s, and for music H it is 0·4 s. Each waveform of the running ACF and IACF is shown in Figures 2–8. The value of τ_{IACC} for telephone ringing has some variance resulting from the movement of the head, because the microphones were fixed at the ear entrances of the real head, but those of other sound sources are almost 0.

The relationship between values of $(2T)_r$ and τ_e is illustrated in Figure 9. Obviously, a linear relation is observed between $\log 2T$ and $\log(\tau_e)_{min}$, and its regression line is approximately expressed by the equation.

$$(2T)_r \approx 0{\cdot}03(\tau_e)_{min} \quad \text{(s)}. \tag{1}$$

This result suggests that the integration interval of human auditory–brain system may be changed by the $(\tau_e)_{min}$ of the sound sources.

Additionally, according to the IACF, it was found that IACC is affected by the kind of sound sources, as shown in Figure 10, but the fan noise of the personal computer was eliminated from this figure because it was recorded in the office and others were recorded in an anechoic chamber. In the case of the key-punch and telephone ringing, the sound is intermittent, so that IACC might be changed. On the other hand, the IACC of the continuous noise of the air conditioner is changed because of its strong low frequency, as shown in Figure 3(d).

Thus, the minimum duration of hearing perception is closely associated with the effective duration of the sound signal.

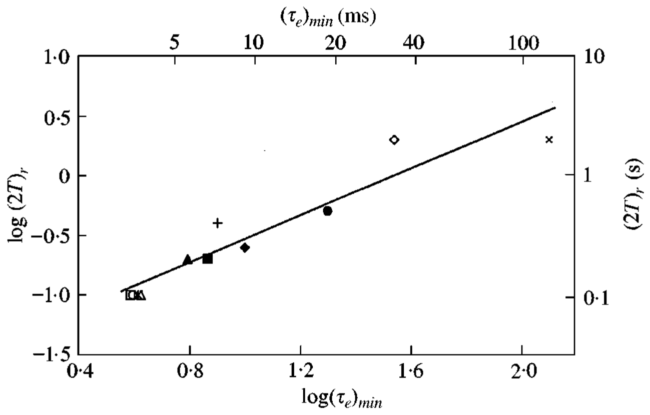

Figure 9. Relationship between $(2T)_r$ and the $(\tau_e)_{min}$ for each sound source: ▲, telephone ringing; ○, the fan noise of the air conditioner; ✳, the sound of the key-punch; △, the fan noise of the personal computer; □, voice; +, music H; ■, music I; ×, motif A; ◇, motif B[5]; ●, aircraft noise A; ◆, aircraft noise B [6].

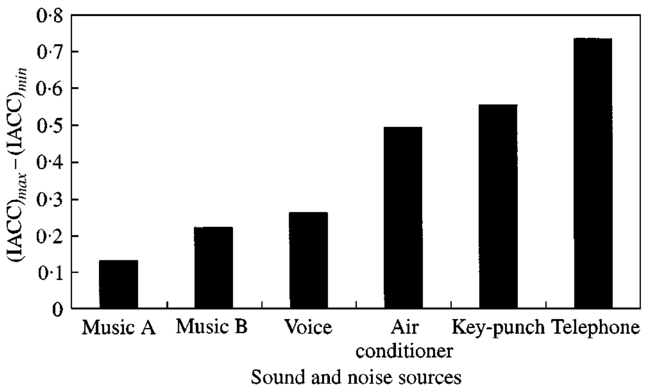

Figure 10. Maximum range of IACC for the noise in an office.

ACKNOWLEDGMENTS

For the analyses of noise source and/or the noise field, we used the new system for identification and evaluation of regional environmental noise that was developed with the support of the JST.

REFERENCES

1. Y. Ando 1998 *Architectural Acoustics, Blending Sound Sources, Sound Fields, and Listeners*. New York: AIP Press/Springer-Verlag.
2. Y. Ando 2001 *Journal of Sound and Vibration* **241**, 3–18. A theory of primary sensations measuring environmental noise.
3. Y. Ando, T. Okano and Y. Takezoe 1989 *Journal of Acoustical Society of America* **86**, 644–649. The running autocorrelation function of different music signals relating to preferred temporal parameters of sound fields.
4. Y. Ando 1983 *Journal of Acoustical Society of America* **74**, 873–887. Calculation of subjective preference at each seat in a concert hall.

5. Y. ANDO 1977 *Journal of Acoustical Society of America* **62**, 1436–1441. Subjective preference in relation to objective parameters of music sound fields with a single echo.
6. K. FUJII, Y. SOETA AND Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 69–78. Acoustical properties of aircraft noise measured by temporal and spatial factors.