

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Contribution of Precisely Apparent Source Width to Auditory Spaciousness

Chiung Yao Chen

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/56616>

1. Introduction

It has been shown that the apparent source width (ASW) for one-third-octave band pass noises signal offers a satisfactory explanation for functions of the inter-aural cross-correlation (IACC) and W_{IACC} , which is defined as the time interval of the inter-aural cross-correlation function within ten percent of the maximum (Sato and Ando, [18]). In this chapter, the binaural criteria of spatial impression in halls will be investigated by comparing with ASW for the auditory purpose assistant to visual attention, which is called source localization. It was proposed that the ASW could properly define directional impression corresponding to the inter-aural time delay (τ_{IACC}) perceived when listening to sound with a sharp peak in the inter-aural cross-correlation function (ICF) with a small value of W_{IACC} . We supposed that the ASW would be sensed not only with regard to the relative amplitudes between reflections in a hall, but the total arrived energies at two ears through the A-weighting network in the brain, termed as listening level (LL) and the temporal characteristics of sound sources. This hypothesis is based on the fact that the spatial experience in a room will be varied by changing the center frequency of one-third-octave band pass noise signal, and the ASW decreases as the main frequency goes up. For the purpose of this chapter, we shall discuss the relationship among some factors, the geometric mean of sound energies at two ears, the reverberation, IACC, τ_{IACC} , and W_{IACC} , and whether they are independently related to the sound source on a horizontal plane. Finally, we have discussed that the ASW impression varied in accordance with the acoustic characteristics of sound intelligibility.

2. Effects of reverberation time and sound source characteristics to auditory localization

2.1. Physical properties of source signals regarding sound localization in a hall

According to the reports by Morimoto [1] regarding the influences of sound localization of spatial perception in a hall, the reverberation energy ($RT60 = 0.3, 0.9$ s) may be treated as the first reflection energy (delay time = 80, 160ms). However, the selection of music is exclusively limited to using Wolfgang Amadeus Mozart's Symphony No. 41, Movement IV as a music source. We intended to prove that the sensitivities on the spatial impression of sound localization will vary depending on the structural characteristics of music. Therefore, the other three sound sources: Motif A (Royal Pavane by Gibbon, $\tau_e = 127$ ms), Motif B (Sinfonietta, Opus 48; IV movement; Allegro con brio by Arnold, $\tau_e = 35$ ms) and Speech (female, $\tau_e = 23$ ms) were adopted. According to the sound field design theory described by Ando [2], the determining factor of an ideal reverberation time length lies in the effective delay of autocorrelation function (τ_e) of sound sources illustrated in Figure 1. The reverberation time of our experiments was set at: short (0.3 s), medium (0.9 s) and long (2.0 s) respectively. The judgments of the apparent sound localization were responded from 12 participants by way of scaling using a normal distribution between two horizontal stimuli angles. The primary analyses of correlations between sound source and auditory localization will presumably the different τ_e proposed by Ando [2]; namely, the significant difference sensation of reverberate image between Motifs will have an influence on human auditory spatial perception of sound sources.

2.2. Analyses of source signals in a hall

The experiences of visual interaction with the direction of sound source at the stage of opera or a classical orchestra have sometimes failed to catch the scene of the performance with respect to the distance or width of the stage. However, it is important and cheering for the audiences to trace and immediately respond to the present player on the stage as if the source directional sensitivity in a diffusing sound field were accurately installed. In this paper, we have tried to compare the source directional sensitivity of spaciousness as caused by early reflections with different azimuth angles. Morimoto [1] reported that of early reflections at the point of subjective equality (it was termed PSE) of spaciousness shows that they are comparable, but early reflection levels seem to be generally slightly lower than the reverberation. That is, the reverberation level correlated well with the early reflections level at the PSE. This means that both energies are fairly proportional to each other and that the average difference is 1.27dB. Barron and Marshall [3] described that the value of lateral energy fraction, as calculated for a series of reflection sequences for two rectangular halls gave virtually identical values no matter whether 80 ms or 100 ms was used as the limiting delay value for the early lateral reflections. Inoue et al. [4] recently reported that the preference of sound impression did not increase with spaciousness throughout, but may have a maximum value at certain spaciousness, that is, the audience does not prefer excessive spaciousness. Hasegawa et al. [5] reported the sound image width was perceived as narrower or wider than the actual presentation region when the sound source width was decreased or increased, respectively by using two loudspeakers were semi-

circularly arranged. Ando [2] reported the most preferred delay time of early reflections after the direct sound differs greatly between the two Motifs. It is found that this corresponds to effective durations (τ_e) of the autocorrelation function (ACF) of source music of 127 ms in Motif A and 35 ms in Motif B. To obtain a degree of similar repetitive features of the sound signals, τ_e values of ACF were analyzed as a phenomenon of stationary random processing (SRP) strictly defined with an infinite length observation (Marple [6]). Concerning SRP for music signal, the estimation of finite length data (2 s) will only obtain an estimation of ACF as Equation (1). As $\tau \ll N$, the estimation of ACF are almost equal to the ACF only in an initial range. Thus, a linear sum of music shows an initial decline of envelope of ACF, and it can be fit to a straight line regression of the power of the normalized ACF (Figure 1). The τ_e values of ACF of music is defined as it crosses to -10 dB to that of delay.

$$\Phi(\tau) = \frac{1}{N} \sum_{n=1}^N x_N(n)x_N(n+\tau) \quad (1)$$

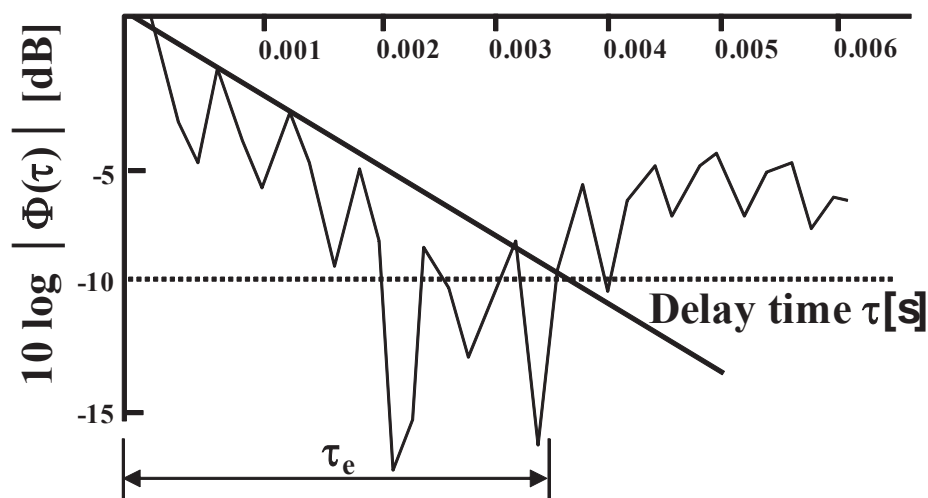


Figure 1. Definition of the effective durations (τ_e) of the autocorrelation function

In order to represent the geometrical size of a similar room, the delay time of subsequent reflection is introduced as $\Delta t_2 = \Delta t_1 + 0.8 \Delta t_1$. In this study, the term “auditory localization” was defined as the detection of sound image edge perceived by the auditory event using two loudspeakers as Hasegawa et al. [5].

2.3. Subjective judgments of sound localization

A method of adjustment using LED unit by the subject was employed in this experiment. The subjects could switch the edge direction carefully with a LED unit equipment (Figure 2), as they were asked to answer the angle of edge direction to the maximum possible under the auditory spaciousness they perceived.

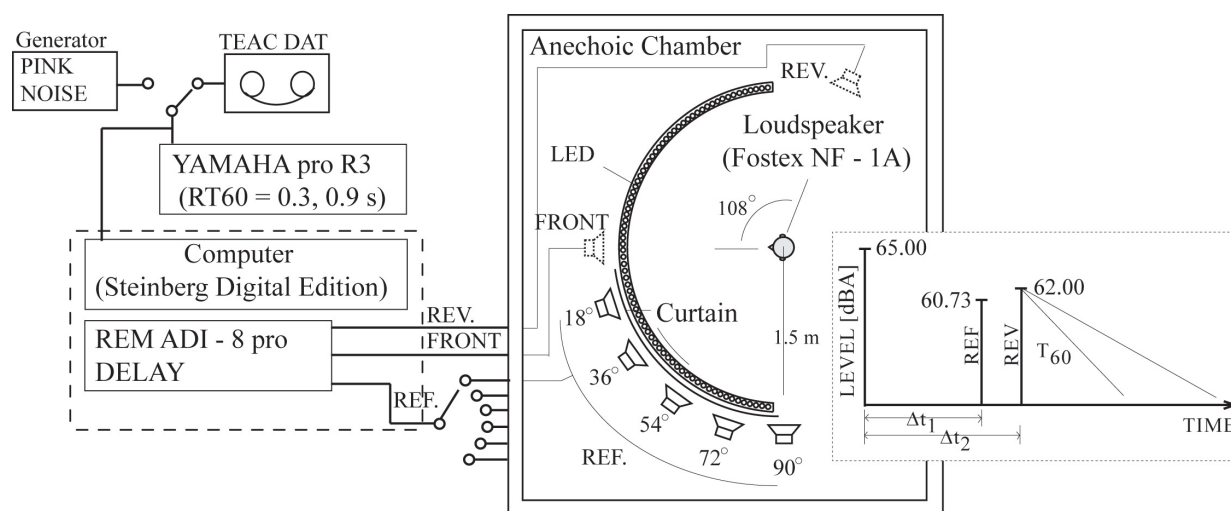


Figure 2. Measuring set-up

1. Apparatus

Figure 2 shows the experimental arrangement. Seven loudspeakers were arranged in the semi-anechoic chamber of the acoustical studio at the Chaoyang University of Technology. The first loudspeaker was in front of the subject at a distance of 1.5m. This loudspeaker was used to radiate the direct sound. One further loudspeaker stood at azimuths of $+108^\circ$, also at a 1.5m distance, used to radiate reverberation. The direct sound was played by digital system controlled on desktop PC derived from a DAT tape recorder (TEAC R-9) and delivered directly to the front loudspeaker. The single early reflection and the reverberant signal with time delay of preferred gap were listed in Table 1. The reverberation time (RT60) was created by a digital reverberator (YAMAHA Pro R3). They were directly delivered to the left horizontal plane by loudspeakers (-18° , -36° , -54° , -72° , -90°) and to the right plane ($+108^\circ$). Mehrgardt and Mellert [7] measured the transfer functions of the ear canal using the impulse response technique from ten directions of the symmetry plane in a free sound field. The peaks of these functions yield about 8% of the different amounts of the shifted curves at these ten directions from 0° to 180° . The curves of 20 subjects overlap closely, if they are shifted along the logarithmic frequency scale. The angles of the early reflection are in five directions of the frontal symmetry horizontal plane (Figure 2). We could simulate five kinds of sound fields, which all consisted of the direct sound plus reverberation and plus early reflection with arbitrary five azimuth angles. The levels of the early reflections and the reverberant signals relative to the direct sounds which were measured by a noise meter (ONO SOKKEI LA-5110) placed above the head of the subject. For the level measurements (SLOW, A weighting, peak), pink noise was used as a source signal. The LED unit could display each 3.0° azimuth angle; the results of these experiments were scaled using normal distribution function as below, the score was 100 as the answer is absolutely right to the present angle, and 0 showed that the answer was a different angle to the present one.

<i>Motif</i>	Δt_1	Δt_2	τ_e	<i>Tempo</i>
A	127 ms	229 ms	127 ms	slowly
B	35 ms	63 ms	335 ms	quickly
S	23 ms	41 ms	227 ms	quickly

Table 1. Experimental arrangements for the three Motifs

Figure 2 simultaneously shows that the level and time delay structure of each signal was constantly arranged for three Motifs respectively for all situations in our experiments. All the data for three Motifs are shown in Table 1.

2. Musical Motif and Subjects

The Motifs used for the experiments were all initial 5s section of Symphony music; they are: (A). Royal Pavane composed by Orlando Gibbons, (B). Sinfonitetta, Opus 48, IV movement composed by Malcolm Arnold, and (S). Speech “In language infuse the T many words become read the small set later.” Poem read by a female, recorded by Burd [8] in the anechoic chamber of BBC. Twelve experienced males, ages 25 ± 2 years, with normal hearing sensitivity served as subjects.

3. Procedures

The subject could switch at will between five azimuth angles using LED unit equipment. After each angle adjustment, the experimenter recorded the results from the LED unit to calculate the score with Equation (2). Reverberation times RT60 of 0.3, 0.9 and 2.0s, and the source signal Motif A, B and S were used for the experimental sound field. The early reflection was radiated at different azimuth angles of -18° , -36° , -54° , -72° , and -90° throughout the three Motifs. Each measurement was repeated three times, yielding a total of 135 experimental results altogether for each subject.

$$SCORE = \frac{1}{\sqrt{2\pi}} e^{-\frac{(angle)^2}{2}} \quad (2)$$

2.4. Analyses of perception on source localization

All data for the twelve subjects are shown together in Figure 3. A three-way (Motif * RT60 * Angle) factor analysis of variance (ANOVA) indicates significant individual difference between three Motifs and five angles ($p < 0.001$, $p < 0.001$) for all experimental conditions. However, the three-way factor analysis of variance indicates less significant difference ($p = 0.029$) between three conditions of RT60. In addition, there is no interference between the three factors for all experimental conditions. This means that all test sound fields could make the subjects perceive spaciousness after the direct sound field no matter what the reverberation time was in the situation of 0.3, 0.9 or 2.0 s. Therefore, the averaged tendency is obvious for three Motifs are obviously higher ($p < 0.001$) as τ_e of ACF of the source signal is longer itself (Figure 4).

Especially, in the case of angle = -54° , scores are quite consistent; the Motifs are clearly independent with the reverberation time. In the case of angle = -36° , the scores were least since subjective diffuseness could be most intense, the source width image was blurred. We conducted a further observation on the measurements of inter-aural cross-correlation coefficient measured by Ando [2] for three Motifs. The measured values of the magnitude of ICF (IACC) for five azimuth angles from -18° to -90° of early reflections are shown in Figure 5. The results of measurements of IACC measured at both ears for music. Especially for Motif A and B, they are noteworthy in connection with the results of source localization in this study.

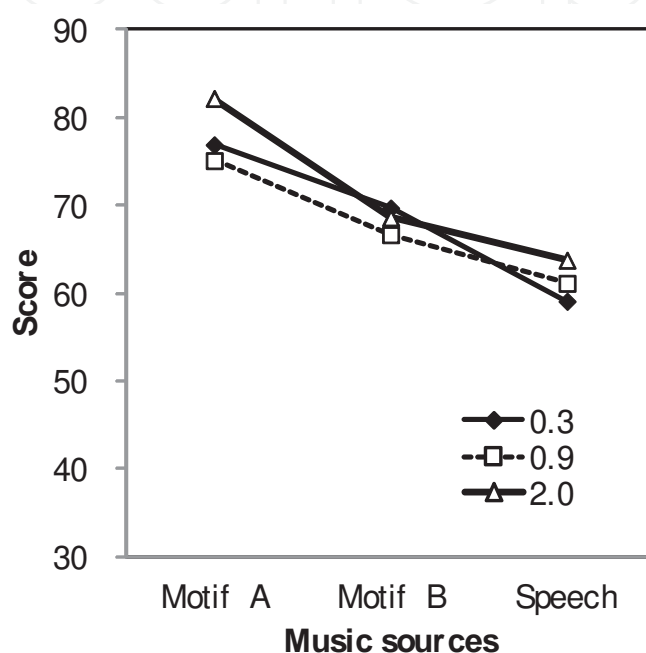


Figure 3. Scores of auditory source directional sensitivity were obtained by changing the coming azimuth angle of early reflection for the three Motifs and different reverberation times. The tendency shows that Motif A obtained the highest accuracy level while speech hit the lowest ($p < 0.01$).

3. Relationship between the envelope of sound image and source characteristics in median plane localization

3.1. Physical properties of apparent source width regarding sound incident angles

To design an indoor sound field, Ando [9] proposed there are three temporal components involved. They are direct sound, first (initial) reflection and subsequent reverberation. This section was further compared with the spatial perception of a media plane in attempt to detect the edge of the sound envelopment composed by such three components. The relationship between source temporal characteristics and apparent source width (ASW) of spatial impression found in above section were reconfirmed, too. The experiment was arranged the direct sound located in front of the subject ($\eta = 0^\circ$, $\xi = 0^\circ$), and the first reflection came from different

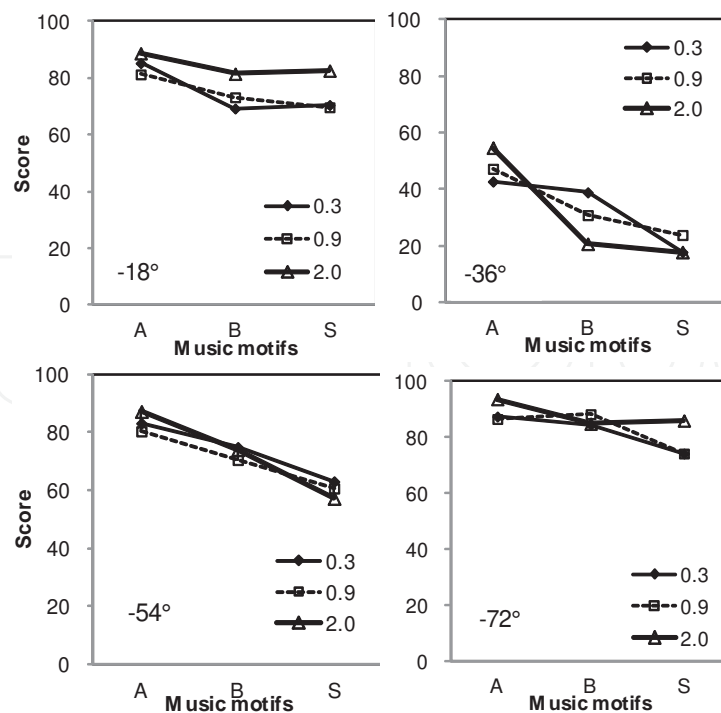


Figure 4. The scores of source width's detection sensitivity function as effective delay of ACF of source in several angles (-18°, -36°, -54°, -72°).

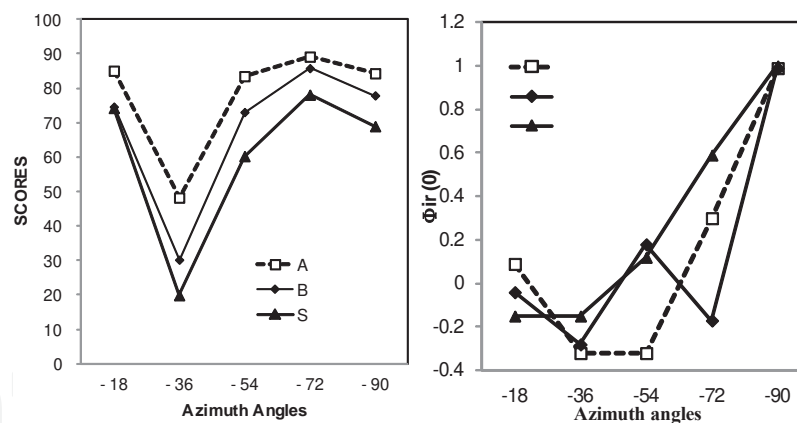


Figure 5. Source directional detection (Left) functions similarly as the tendency of measurements of cross-correlation ($\Phi_{ir}(0)$) (Right) for five azimuth angles from -18° to -90° (contra- clockwise) of early reflections.

vertical angles ($\eta = 18^\circ, 36^\circ, 54^\circ, 72^\circ, 90^\circ$); and reverberation came with energy at a fixed angle ($\xi = 90^\circ$). The subjects were instructed to judge the angles of sound image outline in the sound field by keeping attention on some 5 s duration dry sources of the parts of classic music. The purpose of these arrangements is to confirm that whether subjective judgment of image boundary is affected by reverberation time or not. Secondly, is the ability of edge localization independent with the angles of first reflection in media plain?

3.2. Studies reviews of apparent source width at the median plane

We have experienced in edge detection of the sound image envelope in relation to the localization of sound sources on a horizontal plane in an indoor sound field (Chen [10]). According to several reports by Morimoto ([11, 12] and [13]), they confirm that the localization accuracies almost always depend on the presence of spectral cues of median-plane localization, and that most sound images are recognized by both binaural disparity cues and spectral cues at a certain biased direction. However, Morimoto applied only white-noise through a band-pass filter as a sound source, but not a contribution to the aid of building acoustic design. We referred to the results as Morimoto reported [14] on the energy setup of whole reflections within a horizontal plane for apparent source width (ASW) in a hall, and found that source temporal cues have a strong influence on the edge detection of the sound image envelope using the auto-correlation technology proposed by Ando [9]. The purpose of this study focused on the problem of whether or not the localization tests of source images in the upper hemisphere in a median-plane need both binaural cross-correlation cues and dynamically temporal cues. Temporal cues mean that the spaciousness of a sound field depends upon not only on inter-aural cross-correlation but source characteristics themselves. After all, the coming orientations of initial reflections to the audience in a hall indicate an important design theory which is to be improved by source image creation.

Barron and Marshall [15] identified the arrival time of reflections by 80-100 ms after the direct sound. In terms of Morimoto et al. [16], spatial impression comprises of at least the following two components. One is an auditory source width (ASW) which is defined as the width of the sound image fused temporally and spatially with a direct sound's image and the other is listener envelopment (LEV) which is the degree of the fullness of sound images around the listener, excluding the sound image composing ASW. The auditory spaciousness was inquired under initial reflection and reverberation in a concert hall by Morimoto et al. [16]. The difference limen applied to subjective auditory perception. The sound pressure of direct sound as the standard made that of initial and reverberation noticeable. The point of subjective equality (PSE) applied to identify the least sound pressure level under the timing of just-noticeable difference of direct sound energy. The outcomes show that the listener's auditory spaciousness is not affected by delayed reflections and reverberation time at the sound pressure level (SPL) by 1.27 dB between the two reflections.

Room shape, reverberation time and first delay time are often taken into account in designing an indoor sound field; therein, the sidewall planning influential to reflections is valued in particular. However, the azimuth reflection is overlooked. From the reports of [10, 18], there is a correlation between the apparent source width (ASW) and the direct sound, initial reflection and subsequent reverberation of Motifs of which a sound field comprised might compose varied spaciousness of apparent sound source or edge detection of sound image envelopment. The experiments were conducted after validating and verifying the accuracy of the temporal and spatial components to prevent the spatial split. By Chen [10], the temporal characteristics of music do affect the auditory spaciousness of apparent sound source whereas how reverberation time impact on spaciousness is in need of further verification. The human auditory system is sensitive to sounds at frequencies between 1000-4000 Hz pursuant to an

equal loudness contour. Asahi and Matsuoka [17] failed to explain how human ears discern the frequencies. Morimoto et al. [13] employed white noise as the binaural stimuli by 4800 Hz since the azimuth localization depends on the high-frequency sound source in contrast to the low-frequency one. However, the author finds such statement in need of more verification.

This focus of the study is whether or not the localization tests of the source image in the upper hemisphere (Figure 6) in a median-plane need both binaural cross-correlation cues and dynamically temporal cues. Temporal cues mean that the spaciousness of a sound field depends upon not only inter-aural cross-correlation but source the characteristics themselves.

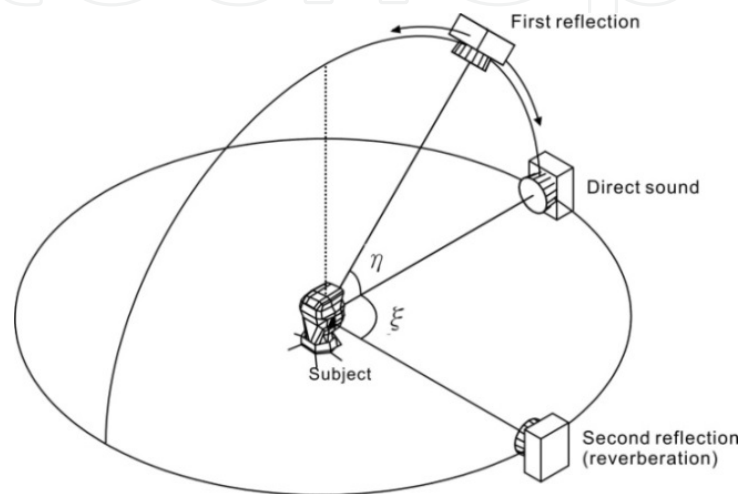


Figure 6. Demonstration of a sound field

3.3. Subjective judgments of source envelope at the median plane

Figure 7 shows how the subject perceived the sound. There were direct sounds in front of the subject ($\xi=0^\circ$) with first reflection at vertical angles ($\eta = 18^\circ, 36^\circ, 54^\circ, 72^\circ, 90^\circ$) and second reflection (reverberation) in front of the subject at 90° ($\xi=90^\circ$).

1. Arrangement

The spaciousness consisted of the three components which involved direct sound, initial reflection and reverberation and was surveyed to identify the degree of edge detection on sound envelopment in the upper hemisphere in a median-plane excluding other unwanted factors. First, the subject reported that the perceived angle seated at a specified chair of a semi-anechoic chamber by a semi-round LED device with intervals by 3° across 60 LED lamps within a radius of 1.5m in order to determine the angles of subjective edge detection on sound envelopment.

2. Parameters

According to Ando [9], the temporal and spatial parameters of a sound field cover sound pressure level (SPL), first reflection, reverberation time and inter-aural cross-correlation coefficient (IACC) by which the parameters of the three components were set up. Figure 7

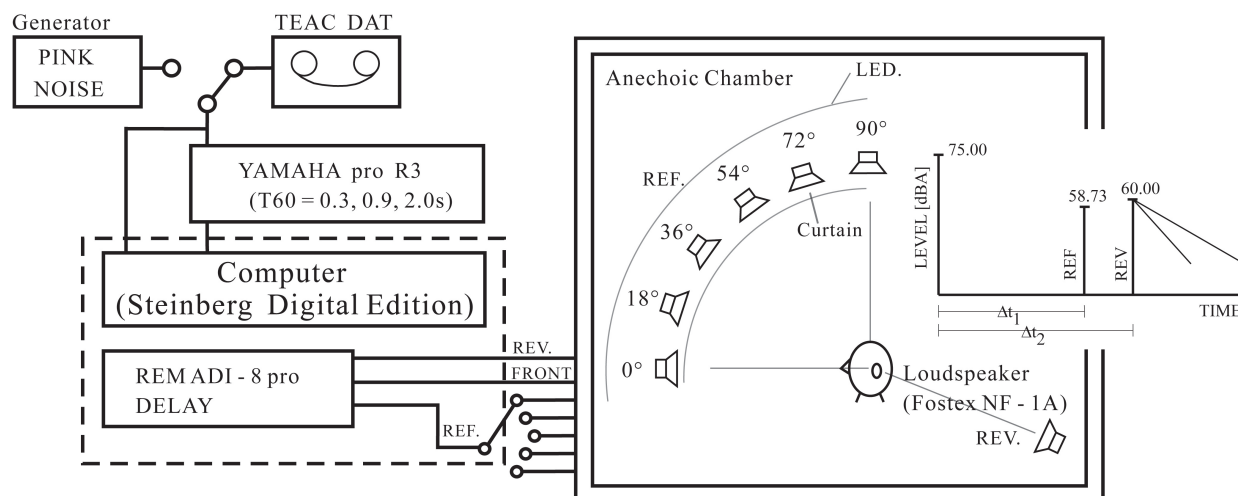


Figure 7. The block diagram of the simulation system for direct sound and two early reflections and the diffused reverberation is attached on the second reflection, which was used in all subjective judgment experiments. Sound pressure levels of the three components were illustrated simultaneously. The direct sound was located in front of subject ($\xi=0^\circ$) with first reflection at the median plane from $\eta = 18^\circ$ to 90° and reverberation at clockwise horizontal plane 90° ($\xi=90^\circ$, $\eta=0^\circ$).

simultaneously shows the setting up of sound energy in compliance with spatial components of sound energy in a common indoor sound field by the SN ratio of direct sound and first reflection by 15 dB and SPL of direct sound and the other two by 75 dB(A) and 60 dB(A). By the report on the auditory perception in a concert hall by Morimoto [8], reverberation can compose a full image of spaciousness as the second reflection with energy more than the first reflection by 1.27 dB. This is the so-called point of subjective equality (PSE). Thus, the energy of early reflections was reduced to 58.73 dB (SLOW, A weighting, peak). Figure 7 shows the equality. The time gap between direct and first reflection sound (Δt_1) was determined pursuant to research by Morimoto [16] under early reflection sound at 50 (ms) and reverberation at 80 (ms) in compliance with the gap by 1.8 times between early and subsequent reflections by Ando [9]. Also, the author arranged the experiments under RT60 = 0.3s (short), 0.9s (medium) and 2.0s (long) to enhance the impact of reverberation time on spaciousness in a sound field.

3. Determination

- Split judgment (Preliminary)

To prevent image split in a sound field, 36 sound fields randomly comprising of the three Motifs (Motifs A-C with time: 5s) under 3 directions of early reflections ($\eta = 18^\circ, 54^\circ, 90^\circ$) and four reverberation times (0.0s, 0.3s, 0.9s, 2.0s) were judged by 15 subjects for 3 times respectively. In this procedure, the subjects confirmed that sound envelopment was perceived as an integrated image without split.

- Edge detection (Primary)

To obtain sound image outline of respective angles, reverberation times and Motifs, 45 sound fields randomly comprising of three Motifs (Motifs A-C with time: 5s) under five directions of

early reflections ($\eta = 18^\circ, 36^\circ, 54^\circ, 72^\circ, 90^\circ$) and 3 reverberation times (0.3s, 0.9s, 2.0s) were judged by subjects through the sensory threshold of adjustment method for three times respectively. In this procedure, the subjects were asked to answer regarding how the location of the edge of sound envelopment was perceived.

4. Subjects and samples

The subjects of two procedures were 15 male students with normal hearing aged 25 ± 2 . In terms of the signal autocorrelation functional theory by Ando [9], a sound source is featured with varied dynamically temporal characteristics critical to spaciousness of a sound field in addition to spectral cues that are called autocorrelation or temporal cues. Table 2 shows details of Motifs A-C.

Source	Title	Composer, writer	Tone	re:ms
Motif A	Royal Pavane	Orlando Gibbons	Andante Downcast	127
Motif B	Sinfonietta, Opus 48; IV movement	Malcolm Arnold	Light Vivid	35
Motif C	Symphony No.102 in B flat major; II movement	Franz J. Haydn	Adagio	65

Table 2. Details of Motifs A-C. Source: BBC (Burd, [8])

3.4. Analyses of subjective source envelope at the horizontal and the median plane

1. Subjective integrity of sound image

The subjective integrity of sound image outline is independent of the angles of first reflection ($\eta = 18^\circ, 54^\circ, 90^\circ$) (three-way ANOVA, $P = 0.900$). Motifs A-C are independent as well (three-way ANOVA, $P=0.322$). Through the ANOVA, subjective integrity is dependent with the reverberation time (three-way ANOVA, $p < 0.001$) and Table 3 shows the results of a Latin Square Design (LSD) analysis of reverberation times. Results indicate that the subjective integrity of the sound image is not affected by the variation of the reverberation time, but both with and without reverberation time.

Means followed by the same letters are not significantly different at 5% level.			
t Grouping	Mean	N	RT60
A	2.9333	45	0.3
A	2.8667	45	2.0
A	2.8444	45	0.9
B	0.5333	45	---

Table 3. LSD of reverberation times

2. First reflection and edge detection on envelopment

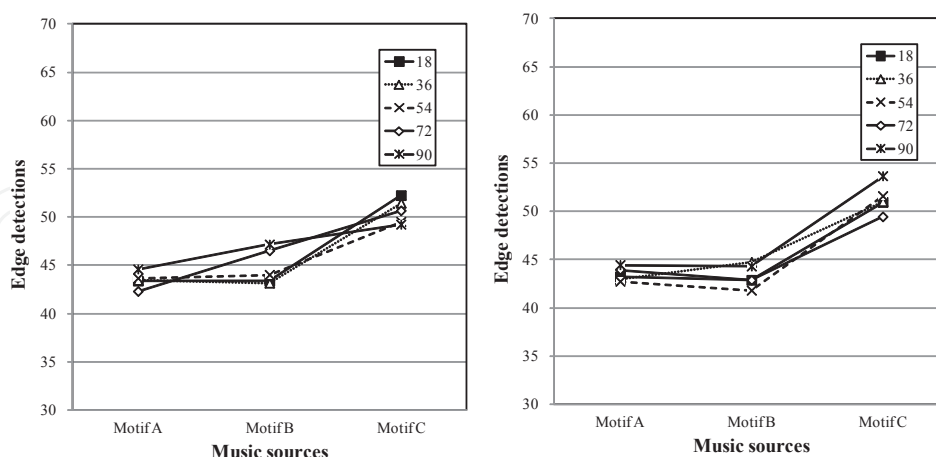


Figure 8. Results of edge detections on Motifs A-C oriented by lateral reflections at the median plane (Left: RT60 = 0.3s ; Right: RT60 = 2.0s)

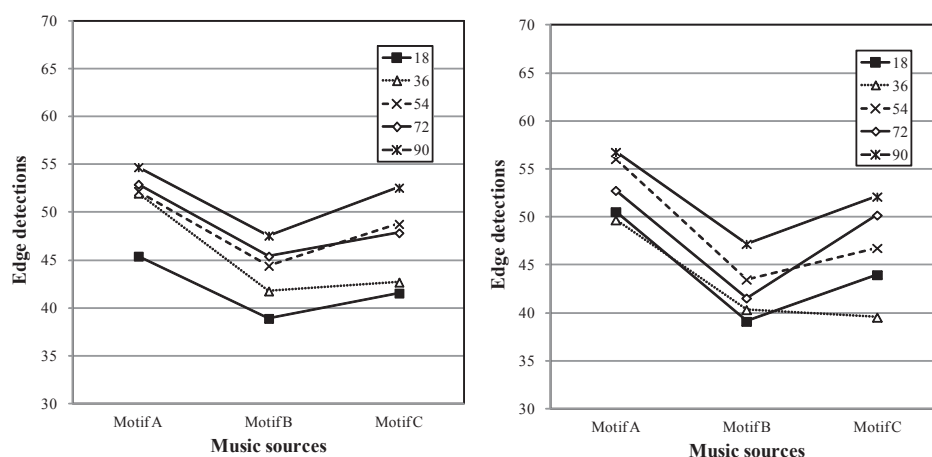


Figure 9. Results of edge detections for Motifs A-C oriented by lateral reflections on the horizontal plane as a reference to Figure 3 (Left: RT60 = 0.3s ; Right: RT60 = 2.0s)

4. Relationship between speech articulation of monosyllable and inter-aural cross-correlation

4.1. An approach on speech intelligibility regarding binaural sensation in a hall

The speech intelligibility for the monosyllables of Chinese in Taiwan area are in agreement with the effective duration of autocorrelation function (τ_e) of the syllable itself in the same reverberation levels were found (Chen and Chan [21]). On the contrary, it was found (Chen [22]) that they are opposite between speech transmission index (STI proposed by Steeneken

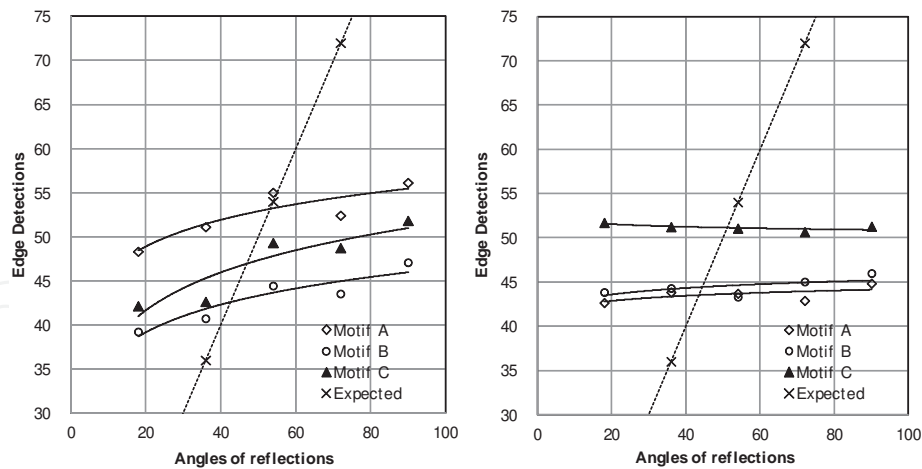


Figure 10. Results of averaged subjective edges values for the significant differences between Motifs A-C oriented by the lateral reflections on the horizontal plane (upper) for mean values at all RT60 conditions, and the source width associated with the τ_e , ACF of the music sources. However, the source width is independent of the reflections on the median plane (see below).

and Houtgast [23]) and magnitude of inter-aural cross-correlation (IACC) where the slope of ceiling were changed in the hall. However, the range of STI (0.5 ~ 0.7) was quite constricted in this study. Takaoka and et al. [24] once used noises and Japanese language to examine the influence of a sound field's reverberation time and IACC (magnitude of inter-aural cross-correlation function) on speech articulation. It was found that under an IACC condition where SN (signal-to-noise ratio) was between -10dB ~ 10dB and reverberation time varied between 0.5s ~ 4.0s, no obvious changes were noticed in speech articulation, and that only when SN was lower than -10dB, IACC affect speech articulation within the range of IACC limited in between 0.5 ~ 1.0. Accordingly, this section focuses on a broadened IACC range (0.34 ~ 0.87), and adopted the paired comparison to identify the relationship between speech articulation and IACC with or without reverberate energy in a hall.

4.2. A generalized theory of biaural measurements in a concert hall

1. The IACC of a sound field

In the field of room acoustics, Ando [9] adopted the magnitude of inter-aural cross-correlation function (IACC) to elucidate human ear's spatial impression on sound field, and also determined main diffuse grades and perception of horizontal directionality of acoustic source in a sound field. Tessier and et al., [25] stated that directionality of acoustic source was a physically front-end mechanism of cocktail effect. They researched on voice articulation in noisy environment through acoustic source separation. But the purpose of study would not feed to the systematical hall design. Ando [9] hypothesized that impulse response of each ear on the path of sound transmission was $h_{nl}(t)$ and $h_{nr}(t)$ respectively. Their inter-aural cross-correlation function can represent human's subject sound localization or spatial impression against sound field. The signals $fl(t)$ and $fr(t)$ of sound's arriving in the ears can serve to express that IACC represents brain's spatial treatment mode, which is defined as follows:

$$\Phi_{lr}(r) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} f'_l(t) f'_r(t + \tau) dt, |\tau| \leq 1ms. \quad (3)$$

Both $f'_l(t) = f_l(t) * S(t)$ and $f'_r(t + \tau_1) = f_r(t) * S(t)$ refer to signals passing through the A-weighting filter which corresponds to hearing perception $S(t)$. Standardized IACC can be modified to Equation (4) from Equation (3) as follows:

$$\varphi_{lr}(\tau) = \frac{\Phi_{lr}(\tau)}{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}} \quad (4)$$

$\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ are monaural autocorrelation functions when delaying τ at the original point (autocorrelation function equals to the average sound intensity of both ears when $\tau = 0$), and total energy arriving both ears is:

$$\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)} \quad (5)$$

However, standardized cross-correlation function in a real room sound field can be modified as follows based on number of reflected sounds and their difference in energy:

$$\varphi_{lr}^{(N)}(\tau) = \frac{\sum_{n=0}^N A n^2 \Phi_{lr}^{(n)}(\tau)}{\sqrt{\sum_{n=0}^N A n^2 \Phi_{ll}^{(n)}(0) \sum_{n=0}^N A n^2 \Phi_{rr}^{(n)}(0)}} \quad (6)$$

where $\Phi_{lr}^{(n)}(\tau)$ is the cross-correlation function forming in both ears by the n^{th} reflected sound; Therefore, the grade of inter-aural cross-correlation function can be defined as Equation (7):

$$IACC = \left| \varphi_{lr}^{(N)}(\tau) \right|_{\max} \quad (7)$$

and the maximum delay of signals between both ears is limited to $|\tau| \leq 1ms$.

Moreover, when point source defuses on plane angle ξ (with the front $\xi = 0$ as datum point) and if the source signal is broadband noise between low and high cut-off frequencies, f_1 and f_2 , the inter-aural cross-correlation function can be modified to:

$$\Phi_{lr}(\tau) = H_{lr} \left[\frac{2}{\Delta\omega(\tau - \tau_\xi)} \right] \sin \left[\frac{\Delta\omega(\tau - \tau_\xi)}{2} \right] \cos \left[\frac{\Delta\omega_c(\tau - \tau_\xi)}{2} \right] \quad (8)$$

where H represents power value of each function, τ_ξ represents the left and right delay caused by horizontal angle ξ , and ω is frequency of filter.

where

$$\Delta\omega_c = 2\pi(f_2 + f_1), \Delta\omega = 2\pi(f_2 - f_1) \quad (9)$$

Figure 11 explains relationship between inter-aural cross-correlation function and various reference factors, while variation width (W_{IACC}) of cross-correlation is as follows when $\frac{\Delta\omega_c}{2}$ is minimal:

$$W_{IACC} \approx \frac{\Delta\omega_c}{4} \cos^{-1} \left(1 - \frac{\delta}{IACC} \right) \quad (10)$$

where δ is the percentage of human ear that can serve to judge change existing in IACC, which is 0.3 normally; Equation (10) shows that maximum W_{IACC} generates the maximum directional perception against acoustic source at horizontal angle ξ . On the contrast, when $IACC < 0.15$, subjective diffuseness can be perceived.

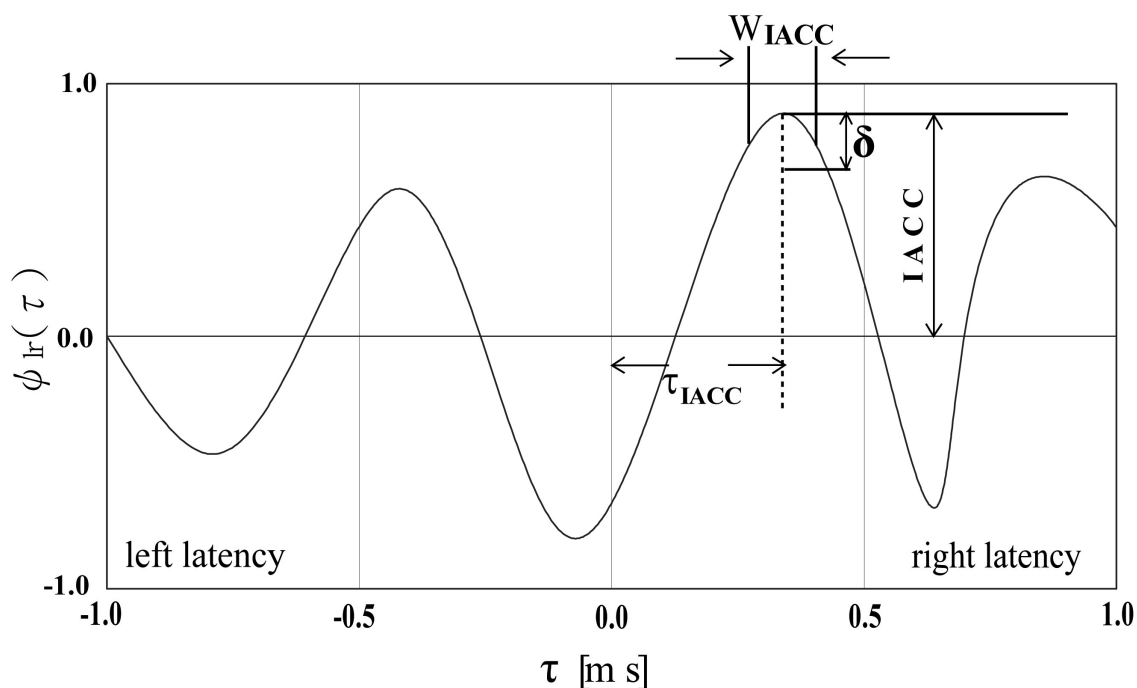


Figure 11. The eigenvalues of standardized IACC can be modified by Equation (4).

Sato, Mori and Ando [26] proposed magnitude of inter-aural cross-correlation function (IACC) and variation width of cross-correlation function can determine magnitude of acoustic sound

width (ASW). Since the source used in the experiment was 1/3-octave noise, they found perception of ASW was lessened when center frequency (125Hz – 2kHz) width was enlarged. Therefore, they proposed to define W_{IACC} as a span during which IACC was within 10% of profile scope of cross-correlation figure's maximum, which corresponds to ASW. Schroeder et al [27] found correlation between $IACC_t$ ($t = 50 \sim 140\text{ms}$) and listening preference. Therefore, IACC indeed increases its applicability to subjective diffuse of sound field. As stated in section 2., Chen and Chang [28] used sound field of two reflected sounds to investigate directional perception of subjective source with musical samples, and he found IACC was the dominating factor and inhibited by magnitude of total reflected sound and length of reverberation. Ohnisi and et al., [29] utilized metro station to research transmitting articulation of sound and found that under influence of 1/3-octave background noise, IACC of the diffuse sound field decreased with increase of sound frequency, and articulation of sound transmission was lowered too. Thus property of spatial sound transmission in sound field is related to variation of IACC.

2. Subjective word intelligibility in sound field

As early as the age when telecommunication devices, such as telephone, were first invented, articulation test has been adopted to test perceptibility of auditory sense against language. Such test was employed to test communicating quality between transmitter and receptor. But now, it is applied to test articulation of telecommunication. Licklider and Kryter [30] conducted objective physic and subjective psychological experiments for speech intelligibility (STI) in Bell Telephone Laboratories and Harvard University's Psychological Sound Laboratory respectively in order to establish a set of effective mono-syllabic test lists, known as Harvard P.B.50 word score (Phonetically Balanced Word List, PB). To expel suggestive factors of other speech voice signals during process of measurement from influencing identical accuracy of STI, articulation test lists were composed of a series of common mono-syllables, with each syllable made up of consonant and vowel. Currently, there are many experimental measure methods which adopt this mono-syllabic speech scale in the world such as Diaz and Velazquez's [31] mono-syllabic speech scale for Spanish. Chen and et al., [32] compiled 108 common vocal samples from New Chinese Phonologic Rhymes, which were used in Taiwan area, and summarized six sets of Chinese mono-syllabic subjective speech articulation scale item (hereinafter refer to as "articulate scale") from them. Based on these 6 mono-syllabic sets, this study found reverberation time (RT60) in room less than 1.5 s in the space of the auditoriums, about $<12000 \text{ m}^3$, the result of STI was consistent with subjective speech articulation and only varied more obviously in few mono-syllables with nasal or voiceless alveolar affricate consonant. To calculate the ability of speech intelligibility, this study calculated percentage of syllable number the subjects could note down accurately during the test to represent correct answer rate and spatial subjective speech intelligibility.

Morimoto, Sato and Kobayashi [33] proposed interaction between word-intelligibility and word-difficulty, where highly intimate words were used to the perceived test sound. In word-intelligibility, the levels of word recognition were the intelligibility percentage of the test sound released to the subject. The experiment result showed that, word-intelligibility and word-difficulty were extremely negatively correlated. Assuming in a sound field with a higher speech transmission index in a public space, the perception of a word-difficulty was higher

than that of word-intelligibility and could be assessed more strictly. When investigating reliability of mono-syllabic speech scale, the issue that Chinese mono-syllables undeniably contains mono-syllables, meaningful and meaningless. This study conducted the subjective psychological experiment by adopting paired-comparison method to solve such vague signals of language expression. By bold assumption that there were only identification method of two-sample which was relatively unaffected by “meaningfulness” and “meaninglessness”, so the subjects could easily identify which one was more intelligible. Similarly, Licklider [34] investigated IACC’s effect on word-intelligibility under noise masking and found that except the effect of SN ratio, decrease of IACC could improve word-intelligibility in the way that mono-syllables were replaced by short sentences. Chen [35] arranged the recordings of mono-syllables in 7 halls, and found that effect both word-difficulty and word-intelligibility could be separated clearly using accumulated cepstrum of the speech voice.

4.3. Subjective attributes of the sound fields with two initial reflections in relation to mono-syllables intelligibilities

1. Setting and configuration of objective physical quantities

Since the variation range for expanding IACC conditions in the experiment of Takaoka and et al., [24] was too narrow, speakers in semi-anechoic chamber were employed to serve sound field simulation of fewer reflection sound energy from various angles. This system was based on the method of IACC simulation design by Damaske and Ando [36], which allowed individual energy and time delay of direct and reflection sounds in sound field. It was equipped with reverberator to feed subsequent reverberant energy so as to decrease quantity of loudspeakers. This study cited the sound field simulation system in the subjective assessment experiment by Damaske and Ando [36] as reference. In order to simulate different circumstances of room IACC’s effect on intelligibility of mono-syllables, this study hypothesized a direct sound in straight front of the subjects, the first and second reflected sounds were hypothesized to transmit to the subjects from different azimuth angles. To further explore the inference by reverberation time of the room, part of the energy of subsequent-RT (RT60) were added to the first and second reflected sounds simultaneously, and then simulated to configure the loudspeakers in the semi-anechoic chamber, whose diagram is shown as Figure 12.

For convenience of the experimental configuration of sound simulated quantity, IACC should be first calculated by adopting Equation (7) from the values of $\Phi_{ir}(\tau)$ and $\Phi_{rr}(\tau)$ measured by Ando [9]. Next, the loudspeakers should be arranged within the range as to generate the IACC in the range of 0.3 to 1.0, where the white noise served as sound source and the dummy head to receive signal. As illustrated in Figure 12, θ_1 and θ_2 were set at 90° and 108° respectively, and with configuration of the IACC measurement was 0.34, 0.56, and 0.87 respectively.

Based on the above simulated configuration, loudspeakers on both sides were added RT energy and set as RT60 = 0.5s and 2.0s respectively. All loudspeakers were 1m from center of the subjects’ heads and 1.2m from the ground, while sound pressure was set as 65 dB (SLOW, A weighting, peak) at upper center of the head. Initial reflected sounds mainly simulated the reflection of right and left walls in the simulation of a hall. The delay time and details of sound field are shown as Table 4.

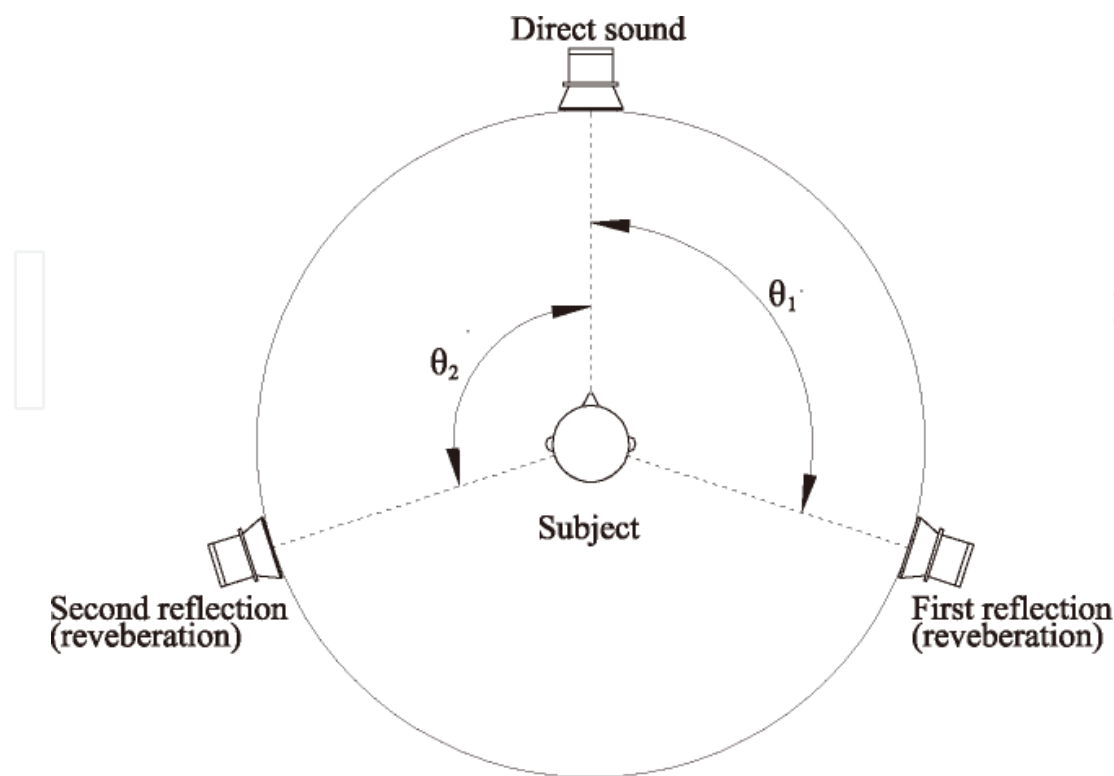


Figure 12. Assumption of IACC configuration was composed by three loudspeakers arranged at different azimuth angles.

2. Sound source

Mono-syllables were same as the research [37] on the correlation between speech intelligibility and continuous brain wave recorded on cerebral cortex, where mono-syllables with higher subjective word-intelligibility such as /heh4/, /ian1/ and /tzuen1/ were figured out, and then compared them with the lower /yu2/.

3. Subjects and experimental method

Total 58 students with average age 23 ± 5 were enrolled as subjects. These subjects were requested to listen and directly answer to experimenter as speech intelligibility. They sat on a fixed chair in the semi-anechoic chamber and concentrated located as Figure 13. The speakers (FOSTEX, NF-1A) were covered with cloth in the semi-anechoic chamber with the light dim. Subjects kept their heads straight ahead and were not allowed to turn, and a repeated test should be avoided in order to avoid over familiarity with the speech samples and thus impairing independence of comparison between sample pairs modified by the assumption of Thurstone's CASE V [38]. This is an obedience to CASE V in paired-comparison theory, that a pair of rivals is independent of each other. In order to quantify the psychological responses of subjective word intelligibility, this study adopted paired-comparison method to gather the scale values of individual syllable, by pairing individual Chinese mono-syllable samples with sound field setting of IACC randomly, and took three different events which had $RT60 = 0.0$ s, 0.5 s, and 2.0 s in turns. Thus each comparison experiment had six samples and 15 pairs, which

were treated by different quantified values would be yielded under different IACC and RT60 settings. In distribution of time in psychological experiment, response time from prompting time was 10 s, while interval of prompting between every two samples was 2 s. Each speech dry source had a span about 0.3 s in average, thus time required by every 15 pairs was 3:15 min. Listening test of each speech had 60 pairs. With four speeches completed total 240 pairs of differentiating pairs which were done in four working days.

Items	Conditions
Azimuth angles	Direct (Ch1, 0 deg. straight front to subjects), 1 st reflection (Ch2, 90 deg. Ch3,108 deg.), 2 nd reflection (Ch4, 90 deg., Ch5, 108 deg.); Added RT energy (Ch2 90 deg., Ch3, 108 deg., Ch4, 90 deg., Ch5, 108 deg.)
Delay gap between the direct and the reflections,	IACC(0.34)- direct : 63.6 dB(A), 1 st reflection: 62.7dB(A) delay (9.46ms), 2 nd reflection: 62.7 dB(A) delay (17.04ms)
and its SPL setting	IACC(0.56)- direct : 62.8 dB(A), 1 st reflection: 60.8dB(A) delay (10.84ms), 2 nd reflection: 48.8 dB(A) delay (19.51ms) IACC(0.87)- direct : 64.6 dB(A), 1 st reflection: 53.4dB(A) delay (15.48ms), 2 nd reflection: 53.4 dB(A) delay (27.87ms)
Reverberation time (RT60)	0.0 s , 0.5 s, 2.0 s
IACC, measured	0.34, 0.56, 0.87

Table 4. Experimental settings

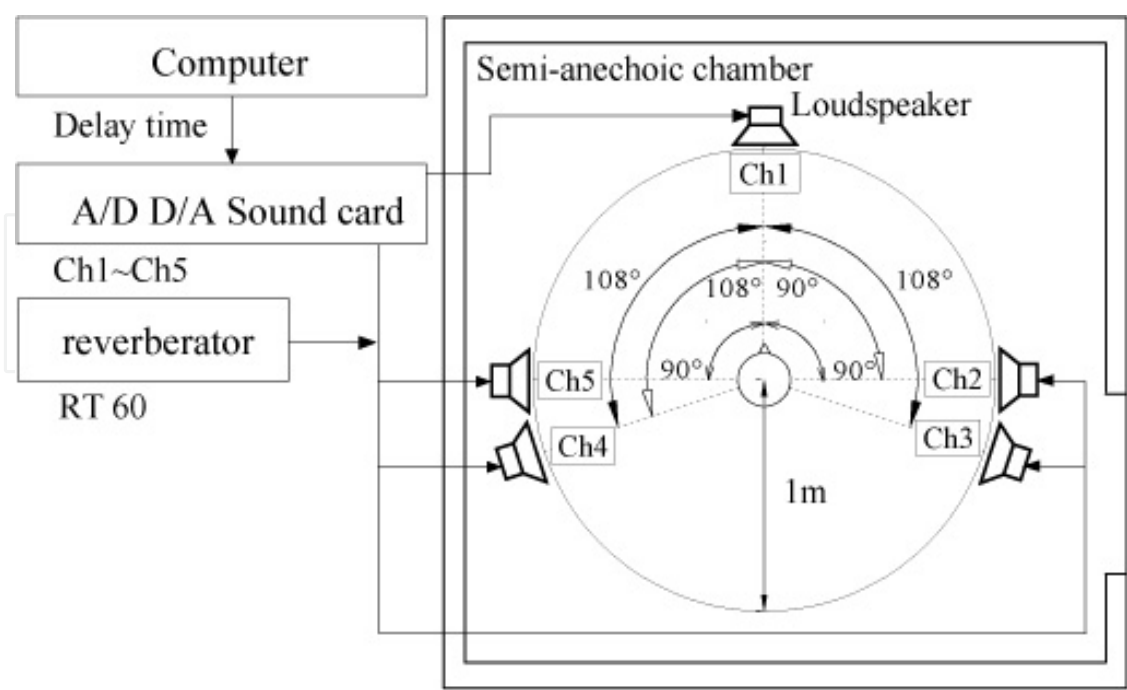


Figure 13. Diagram of experiments

4.4. Analyses of mono-syllabic word-intelligibility

1. The effect of IACC on mono-syllabic word-intelligibility

In order to enhance reliability of the integral answers conducted by paired-comparison method, we counted the numbers of circular-triad once for every subject based on Thurstone’s [38] response consistency test for the experiment of every 15 pairs, through which paired - comparison of these 15 pairs were determined effective questionnaires. Subsequently, a test of goodness of fit for comparison quantification model was performed to verify the scale values met the hypothesis of paired-comparison CASE V by Thurstone [38] with respect to effectiveness of difference between stimuli samples and sample size (Mosteller, [39]).

Based on paired-comparison method CASE V by Thurstone [38], average quantified scale value of word-intelligibility of 58 subjects under the conditions of additional RT60 were calculated and shown in Figure 14 ~ 17. Quantified scale value of subjective word intelligibility of mono-syllables under variation of IACC, 0.34, 0.56, and 0.87 showed that trend of subjective higher word-intelligibility before addition RT60 was significant ($p<0.001$).

By ANOVA, the effect of IACC and RT60 on quantified scale values of mono-syllabic subjective word-intelligibility showed that there exist no interaction between these two factors, two-way ANOVA, $F = 0.27$ and $p = 0.90$. But in the case of an individual factor’s effect on quantified scale values of mono-syllabic subjective word intelligibility, only RT60 presented significantly, two-way ANOVA, $F = 96.38$ and $p < 0.001$), while the effect of IACC had lower significance, two-way ANOVA, $F = 5.34$ and $p < 0.05$. This result reconfirm that RT60 is independent of IACC in sound field, no matter when with regard to musical preference (Ando [9]) or word-intelligibility.

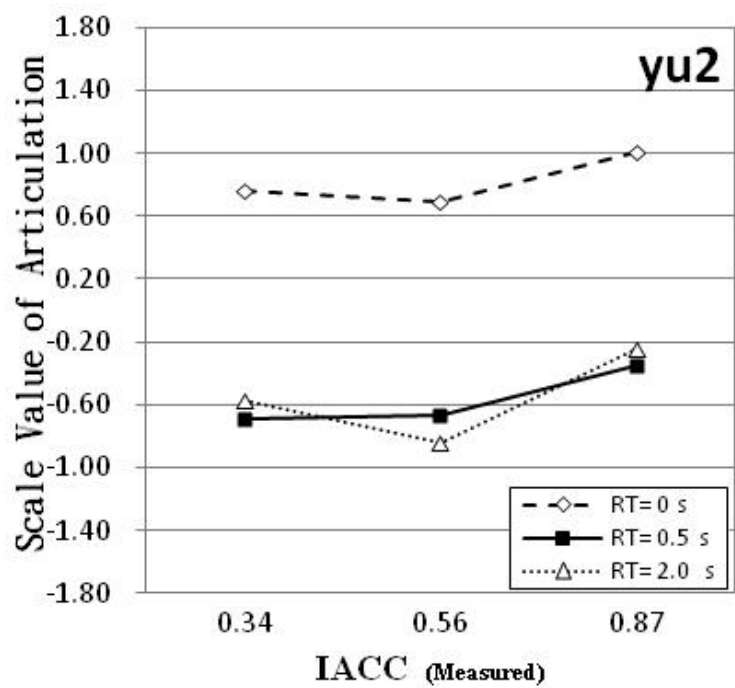


Figure 14. Results of syllable “Yu2”

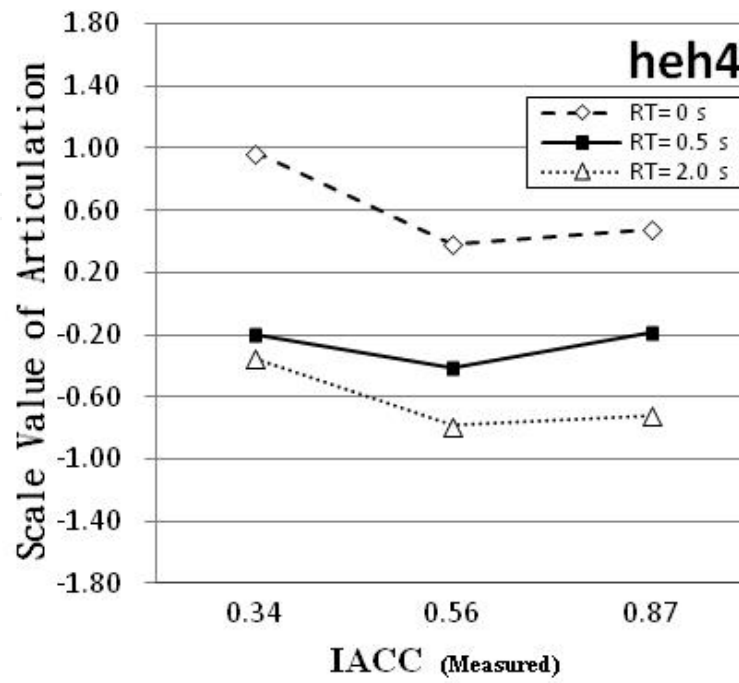


Figure 15. Results of syllable "Heh4"

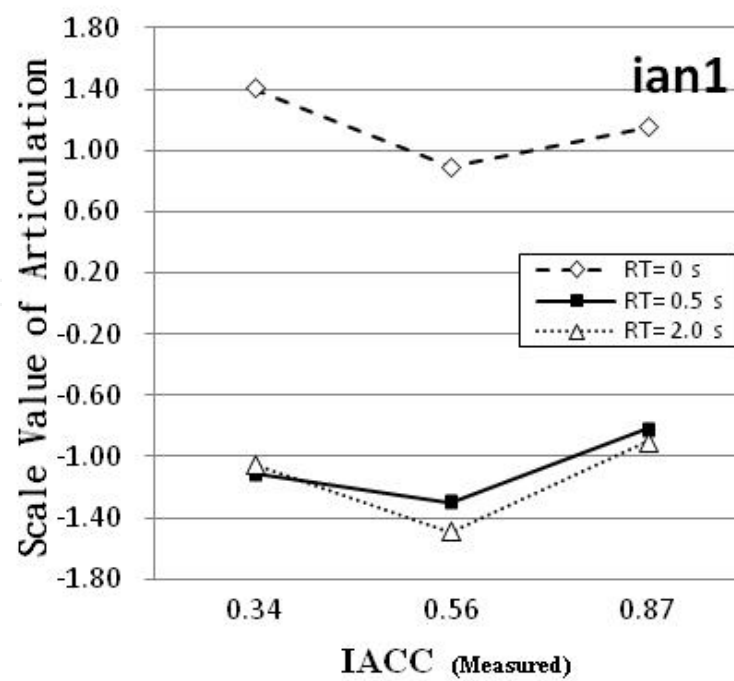


Figure 16. Results of syllable "Ian1"

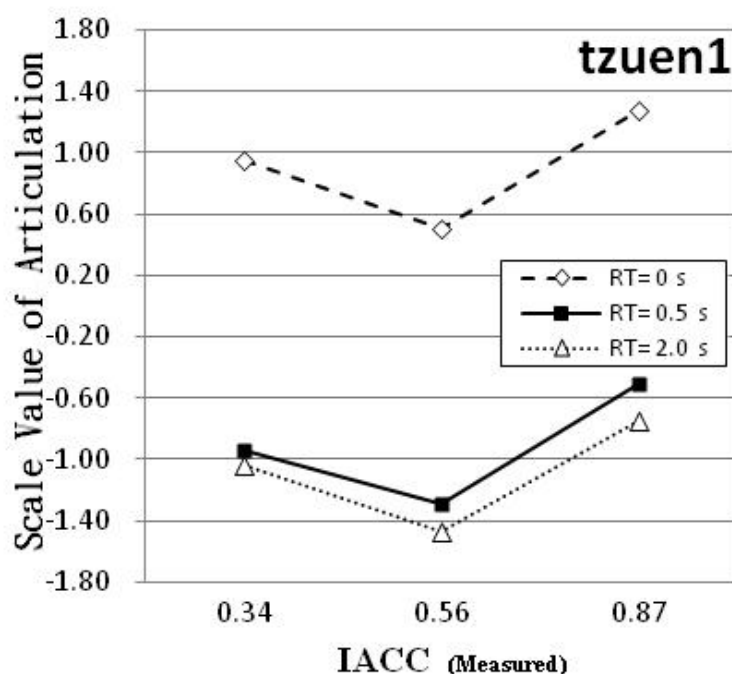


Figure 17. Results of syllable “Tzuen1”

In investigation of the effect of RT60 along on quantified scale values of mono-syllabic subjective word-intelligibility with the setting RT60 = 0.0 s, 0.5 s, and 2.0 s, more significant effect of IACC's variation did not presented. Thus only one-way ANOVA analysis under the environment with RT60 existence and not existence could be performed. The result showed that the effect of IACC's variation was significant in the environment with RT60, by one-way ANOVA $F = 3.74$ and $p < 0.05$. It was doubted of the faith of the results on word-intelligibility is usually changed with regard to IACC in the circumstance of only SN was lower than -10 dB found by Takaoka and et al., [24]. We identify that two reflections of the sound field were not harmful for the word-intelligibility in our settings, and there was no background noise employed here. The setting of RT60 = 0.5 s and 2.0 s adopted here is 1.27 dB in relation to the reflections without reverberant energy at the PSE as stated above (section 2.). Therefore, reflection with RT60 will enhance the variation of IACC on word-intelligibility.

2. The effect of RT60 on quantified scale value of mono-syllabic subjective word-intelligibility

It is clear in Figure 14 ~ 17 that quantified scale values of mono-syllabic subjective word intelligibility obviously changes with RT60. Such change is especially significant between RT60 = 0.0 s and RT60 = 0.5 s. In order to figure out difference among them, this study adopted p value of matrix of Fisher LSD method (Table 5) by multiple mean comparison and found that there was significant difference in quantified scale values of word-intelligibility between RT60 = 0.0 s and RT60 = 0.5 s, $p < 0.001$, while there was no significant difference between RT60 = 0.5 s and RT60 = 2.0 s, $p = 0.297 > 0.05$. This result is similar to that of ANOVA on quantified scale values stated as above, suggesting variation between environments of word-intelligibility with

and without RT60 was significant. Therefore, Takaoka et al. [24] investigated the cross effect of RT60s in sound field on grades of IACC and found that word-intelligibility between 0.5 s and 4.0 s corresponded with the conclusion that grades of IACC were independent from each other. This study complemented the phenomenon that quantified scale values of subjective word intelligibility was influenced by grades of IACC.

Similarly, by testing p value in the matrix of Fisher LSD method (Table 6) with multiple mean comparison it was clear that there was significant difference between quantified scale values of word-intelligibility of IACC(0.34) and that of IACC(0.56), $p = 0.025 < 0.05$; there was also significant difference between that of IACC(0.56) and that of IACC(0.87), $p = 0.004 < 0.05$; while there was no significant difference between that of IACC(0.34) and that of IACC(0.87), $p = 0.445 > 0.05$. Therefore, it was clear from multiple mean comparison test that the effect of variation in IACC on mono-syllabic word-intelligibility was similar to the variation of musical preference in sound field, which were both related to magnitude of data of standardized IACC grades (Equation (4)). However musical preference was inversely proportional to that and was here inversely proportional to mono-syllabic word-intelligibility, by one-way ANOVA $F = 3.74$ and $p < 0.05$. This finding reconfirms that word-intelligibility under varied IACC is associated with nonlinear response in evaluating the subjective localization of sound sources studied above (Figure 5 of section 2.).

LSD test; variable; Probabilities for Post Hoc Tests. Error: Between MS = .11403, df = 27.00			
RT60	{1} 0.872	{2} -0.706	{3} -0.853
0.0 s	—	0.000*	0.000*
0.5 s	0.000*	—	0.297
2.0 s	0.000*	0.297	—

Table 5. The results of RT60 effect evaluated using p value of matrix of Fisher LSD method

LSD test; variable; Probabilities for Post Hoc Tests. Error: Between MS = .11403, df = 27.00			
IACC	{1} -0.155	{2} -0.481	{3} -0.049
0.34		0.025*	0.445
0.56	0.025*		0.004*
0.87	0.445	0.004*	

Table 6. The results of IACC effect evaluated using p value of matrix of Fisher LSD method

3. Relationship between the parameters within wave's characteristics of IACC and word intelligibility

In order to figure out the correlation between IACC and mono-syllabic word intelligibility in detail, this study used dummy head measurement system to detect parameters which were

grades of standardized IACC, delay of inter-aural cross-correlation function (τ_{IACC}), and width of the inter-aural cross-correlation function (W_{IACC}) (Table 7). Sato, Mori and Ando [26] stated in their research that IACC and W_{IACC} could determine acoustic source width (ASW). According to Table 7, the measured data of W_{IACC} in this study was not correlated well to IACC, while τ_{IACC} and IACC showed the opposite trend. Of course, its effect on mono-syllabic word intelligibility also presented RT60 condition under RT60 = 0.5s and 2.0s.

IACC	0.34	0.56	0.87
τ_{IACC}	0.22	0.06	0.09
W_{IACC}	0.19	0.18	0.18

Table 7. The parameters are picked up by wave's characteristic of IACC

5. Conclusions

These facts of section 2. and 3. point out that the temporal characteristics of source signal should be taken into account when estimating and measuring physical measurements, like the lateral energy fraction and the inter-aural cross-correlation coefficient, to estimate source localization sensitivity. For section 4., the experiment of judgment through paired-comparison method, quantified scale values of word-intelligibility was generated based on the hypothesis of CASE V cited by Thurstone [38]. The results show that existence of reverberant energy in a sound field had effect of mono-syllabic word-intelligibility, and that variation of IACC did too. Four mono-syllables with different word-difficulty, subjective mono-syllabic word-intelligibility had certain similar reaction trend under conditions of different IACC and RT60. Results of inductive statistical analyses are shown as follows:

1. As shown in Figure 3, reverberation does not suppress the degree of source directional sensitivity as early reflections after the direct sound, if their ratios of lateral to frontal sound energy are the same. Even though music source directional concept of auditory distinction is inverse to spaciousness of a sound field. The spaciousness is not at all suppressed by levels of early reflections at the PSE at echo threshold for all levels of reverberation whenever the reverberation (RT60) was fixed at 0.3 or 0.9 s concluded by Morimoto [1] as well.
2. As shown in Figure 4, the source directional sensitivity caused by different source signals is suppressed by τ_e of ACF of itself even if the sound field includes both early reflections and reverberation and with their preferred initial time gap after direct sound signals. This finding is an important problem with which to perceive the localization of performers for assisting visual enjoyment in concert halls. The temporal structure of source signal to auditory spaciousness is first discussed out of sound energy or directional mentioned before.

3. The source directional sensitivity are quicker as the coming direction of early reflection sounds located at the azimuth angle from -36° to -54° (Figure 5) as the early reflection functions as lateral energy fraction in a simulated diffuse sound field. The sound incidence angle of -54° is found upon the deep notch and peak at 54° of the curve in the transfer function of the ear canal entrance in a free sound field, especially in the frequency range from 2 to 4 kHz (Mehrgardt and Mellert [7]). It is obvious that source localization at a horizontal plane angle is dependent upon the transfer function of the ear canal.
4. As shown in Figure 7, with a fixed gap between the sound pressure levels of the three spatial components, direct sound, first reflection and subsequent reverberation, the reverberation discerned will affect the capability of an integrated image envelopment without split, demonstrating that reverberation is crucial factor to the envelopment perceived but the edge judgment of image boundary is not affected by reverberation time (Figure 7). This finding is in harmony with the result of sensitivities on reflective signal localization researched in section 2. The reverberation does not suppress the orientation of both source image edges and reflection incidences in addition to the perception of source image split.
5. As shown in Figure 8, the first reflection from the upper hemisphere at the angles $\eta = 18^\circ, 36^\circ, 54^\circ, 72^\circ, 90^\circ$ does not affect the edge judgment of image boundary for music Motifs A-C. The ability of edge localization is independent with the angles of first reflection in median plane but sound source. Rakerd, Hartmann and McCaskey [19] that found listeners failed to identify noises with roved the location when the spectral structure was at a high frequency because the spectral structure was confused with the spectral variations caused by different location. Such is the fact that music with temporal variation leads to confusion regarding the edge of the sound image with a reflection incidence on the median plane in a diffuse sound field. Morimoto and Nomachi [11] have both explained that localization accuracies of sound images on the median plane produced by both binaural disparity cues and frequency cues. Morimoto, Yairi, Iida and Itoh [20] concluded when the source is a wide-band signal, only higher frequency components (> 2 kHz) are dominant on the median plane localization. However, they did not consider that a source with a wide-band sound in temporal variation provides the changing of the source width conception during a concert. Thus, it is presumably difficult to account for the different locations on the median plane of a music source in a hall except for during a recital of an instrument with a higher frequency tones.
6. As shown in Figure 9 and Figure 10, the difference of Motifs and the subjective judgment of edge detections of sound image outline on horizontal plane are interdependent, and the tempo of music proposed by Ando [9] are related well. This evidences that the temporal cues are important to the subjective edge determination and source localization.
7. Depending on one-way ANOVA for the environment with and without reverberation, the result of word intelligibility showed that variation of IACC (0.34 ~ 0.87) had significant effect on the environment with reverberation (0.5s ~ 2.0s), $F=3.74$ and $p<0.05$. Takaoka and et al., [24] reported that IACC influences on speech articulation within the range of 0.5~1.0 only when SN was lower than -10dB under $RT60 = 0.5s \sim 4.0s$. There is no conflict between

these two results because word-intelligibility was not affected by RT60 varied from 0.5s to 2.0s in our research when reverberation was constantly 1.27 dB higher than the reflections. Reflections with RT60 enhance the variation of IACC on word-intelligibility at the PSE of equal spatial impression in the source width. They have obviously confirmed evidence by similar W_{IACC} of varied IACC's environments in Table 7, which may indicate the source width of sound signal stated above.

- 8. Figures 14 ~ 17 illustrate the interaction between RT60 and mono-syllabic word articulation, which show that IACC's effect on mono-syllabic word- intelligibility significantly varied with span of RT60 ($p < 0.001$ ANOVA).
- 9. Test on matrix of Fisher LSD with multiple mean comparison confirmed in Table 5 showed that quantified psychological scale values of word-intelligibility were significantly different between $RT60 = 0.0$ s and $RT60 = 0.5$ s, $p < 0.001$, while not significantly different between $RT60 = 0.5$ s and $RT60 = 2.0$ s, $p = 0.297 > 0.05$. This finding indicates that the source signal image was buried by reverberation and would defect word-intelligibility such as source split as induced by with or without reverberation as investigated in section 2. Similarly, Table 6 confirmed that quantified psychological scale values of word-intelligibility were significantly different at IACC(0.34) and IACC(0.56), with $p = 0.025 < 0.05$, was significantly different at IACC(0.56) and IACC(0.87) too, with $p = 0.004 < 0.05$, while was not significantly different at IACC(0.34) and IACC(0.87), with $p = 0.445 > 0.05$. The nonlinear responses in evaluating word-intelligibility, source edge and localization of spatial impression at the horizontal plane under varied IACC are presumably influenced by transfer functions of the ear canal entrance as measured by Mehrgardt and Mellert [7].

Glossary of symbols

ASW	apparent source width
IACC	inter-aural cross-correlation
τ_{IACC}	inter-aural time delay at cross-correlation function
ICF	inter-aural cross-correlation function
W_{IACC}	inter-aural variative width at cross-correlation function
LL	listening level
RT60	reverberation time
τ_e	effective delay of autocorrelation function
ACF	autocorrelation function
PSE	point of subjective equality

SRP	stationary random processing
DAT	digital auditory tape cassette
ϕ/r	binaural normalized cross- correlation function
$\Phi_{rr}(\tau)$	mono- aural autocorrelation function
$\Phi/r(\tau)$	binaural cross correlation function
η	vertical angles at an median plane, 0° started from the front of head at ear height
ξ	angles at clockwise horizontal plane, 0° started from the front of head at ear height
LEV	listener envelopment
SPL	sound pressure level
SN	logarithm of signal over noise energy, denotes by decibel
Δt_1	delay gap between direct and first reflection in a defuse sound field
LSD	Latin Square Design
STI	speech transmission index
δ	percentage at the peak of wave form in inter-aural cross-correlation function, as the definition of W_{IACC}
$IACC_t$	time gap of sound signal in inter-aural cross-correlation
PB	Phonetically Balanced Word List
/yu2/	example of a mono-syllable in Taiwanese's life speech

Author details

Chiung Yao Chen*

Architecture Department, Chaoyang University of Technology, Taiwan

References

- [1] Morimoto M, Posselt C. "Contribution of reverberation to auditory spaciousness in concert halls [J]," J Acoust Soc Jpn, (E)10, 2: 1989, 87-92.
- [2] Ando Y, "Calculation of subjective preference at each seat in a concert hall [J]," J Acoust Soc Am, 74: 1983, 873-887.

- [3] Barron M, Marshall A H. "Spatial impression due to early lateral resection in concert hall: the deviation of a physical measure [J]," J Sound and Vibration, 77 (2): 1981, 211-232.
- [4] Inoue T, Nishi T, Wakuri T, Shimizu Y, Kawakami F. "Relation between the lateral component of early reflections sound energy and the spatial impression," Proc Spring Meet Acoust Soc Jpn, 1987: 557-558 (in Japanese).
- [5] Hasegawa H, Takehashi K, Ayama M, Kasuga M. "Effects of visual information on sound image localization [J]," J. Image info and Tele Eng, 55, 3: 2001, 455-462 (in Japanese).
- [6] Marple S. L. "A new autoregressive spectrum analysis algorithm," IEEE Trans. Acoust., Speech, Signal Process, ASSP. 28, 1980, 441-454.
- [7] Mehrgardt S. and Mellert V., "Transformation characteristics of the external human ear," J. Acoust. Soc. Am., 61, 1977, 1567-1576.
- [8] Burd V. A. N. "Nachhallfreie Musik fur Akustische Modelluntersuchungen," Rundfunktech. Mitteilungen, 13, (1969) 200-201.
- [9] Ando Y. "Concert Hall Acoustics," Berlin Heidelberg, New York, 1985.
- [10] Chen C. Y. "Effects of reverberation time and sound source characteristic to auditory localization in an indoor sound field," ICSV Cairns Australia, 2007, 9-12.
- [11] Morimoto M. and Nomachi K., "Binaural disparity cues in median-plane localization." J. Acoust. Soc. Jpn. (E) 3,2, 1982, 99-103.
- [12] Morimoto M. and Aokata H., "Localization cues of sound sources in the upper hemisphere." J. Acoust. Soc. Jpn.(E) 5,3, 1984, 165-173.
- [13] Morimoto M., Yoshimura K., Kazhiro I. and Motokuni I., "The role of low frequency components in median plane localization." J. Acoust. Soc. Jpn.(E) 24, 2, 2003, 76-82.
- [14] Morimoto M. and Iida K. "Relation between auditory source width and the law of the first wave front." J Acoust Soc Jpn, 49, 2: 1993, 43-55 (in Japanese).
- [15] Barron M, Marshall A H. "Spatial impression due to early lateral resection in concert hall: the deviation of a physical measure." J Sound and Vibration,,77 (2), 1981, 211-232.
- [16] Morimoto M. and Posselt C. "Contribution of reverberation to auditory spaciousness in concert halls", J. Acoust. Soc., 1989, 87-91.
- [17] Asahi N. and Matsuoka S. "Effect of the sound anti resonance by pinna on median plane localization - Localization of sound signal passed dip filter-," Tech. Rep. Hear. Acoust. Soc. Jpn., 1977, H-40-1.

- [18] Sato S. and Ando Y. "Apparent source width (ASW) of complex noises in relation to the interaural cross-correlation," *Journal of Temporal in Architecture and the Environment*, 2002, 29-32.
- [19] Rakerd B., Hartmann W. M. and McCaskey T. L. "Identification and localization of sound sources in the median sagittal plane," *J. Acoust. Soc.*, 106 (5), 1999, 2812-2820.
- [20] Morimoto M, Yairi M, Iida K, Itoh M, "The role of low frequency components in median plane localization [J]," *J Acoust Sci & Tech.* 24, 2: 2003, 76-82.
- [21] Chen, C. Y. and Chan M. H., "A Study of the Chinese Speech Intelligibility in Halls in Relation to the Autocorrelation Function Model: The Case of Chinese in Taiwan," *J. Archi.*, Architectural Institute of the R. O. C, 57, 2006, 55-68.
- [22] Chen, C. Y. "Investigation on speech intelligibility in respective to the variation of diffusing panels in a hall," 16th Proceeding of the Archi. Institute of the R. O. C, E66, 2004, 336 - 341. (in Chinese)
- [23] Steeneken H. J. M. and Houtgast T. "The Modulation Transfer Function in Room Coustics as A Predictor of Speech Intelligibility, " *Acoustica* 28, 1973, 66-73.
- [24] Takaoka T., Morimoto M., Sato, H. and Semba Y. "Effects of inter-aural cross-correlation of speech intelligibility and background noise on listening difficulty," *J. Acoust. Soc. Jpn.*, NO. 9, 63, 2007, 520-528 (in Japanese).
- [25] Tessier E., Berthommier F., Glotin H. and Choi S. "A case front- end using the localization cue for segregation and then cocktail-party speech recognition," *Proc. IEEE Int. Conference on speech Process (ICsP)*, Seoul (1999).
- [26] Sato S., Mori Y. and Ando Y. "The subjective evaluation of source location on the stage by listeners," In: *Music and concert hall acoustics*. Y. Ando and D. Noson (eds.). Academic Press, London, 1997, 117-123.
- [27] Schroeder M. R., Gottlob D., and Siebrasse K.F. "Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters," *J. Acoust. Soc. Am*, 56: 1974, 1195-1201.
- [28] Chen C. Y. and Chang Y. R. "A Study of Test Method for Absorption Coefficient of Material through Cross-Correlation in an Anechoic Chamber- The porous plane materials as example", *Journal of Technology*, 20 (2005). (in Chinese)
- [29] Ohnisi Y., Maeda K., Morimoto M. and Sato H."Acoustic characteristics of background noise at subway stations,"*Proc. WESPAC IX* (2006).
- [30] Licklider J. C. R. and Kryter K. D. "Articulation tests of standard and modified interphones conducted during flight at 5000 and 35,000 feet (OSRD Report 1976)", Cambridge, MA: Harvard University, Psycho-Acoustic Laboratory (1944).
- [31] Diaz C. and Velazquez C. "A Live Evaluation of the RASTI -Method", *Applied Acoustics*, 46 1995, 363-372.

- [32] Chen C. Y., Chen L. S., and Lin W., "A Study on Evaluation Method of Chinese Articulation Standard of Speech Intelligibility for Sound Field in Taiwan", *Journal of Architecture*, No.43, Architectural Institute of the Republic of China, 2002, 27-36. (in Chinese)
- [33] Morimoto M., Sato, H. and Kobayashi M. "Listening difficulty as a subjective measure for evaluation of speech transmission performance in public spaces," *J. Acoust. Soc. Am.*, 116, 2004, 1607-1613.
- [34] Licklider J. C. R., "The influence of inter-aural phase relations upon the masking of speech by white noise," *J. Acoust. Soc. Am.* 20, 1948, 150-159.
- [35] Chen C. Y., "Syllables Intelligibility in Relation to the Autocorrelation and Cepstrum Model: The Case of Chinese in Taiwan," POMA Volume 15, pp. 015002 (June 2012); Acoustical Society of America, ISSN 1939-800X (online).
- [36] Damaske P. and Ando Y. "Interaural cross-correlation for multichannel loudspeaker reproduction," *Acoustica*, 27, 1972, 232-238.
- [37] Chen C. Y. and Chen U. S., "Cortical continuous brain waves in relation to the speech intelligibility of mono-syllables in Taiwan," master dissertation, Graduated School of Architecture and Urban Planning, Chaoyang Univ. of Tech., Taichung, Taiwan (2010).(in Chinese)
- [38] Thurstone L. L. "A Law of Comparative Judgment", *Psychol. Rev.*, 34, 1927, 273-289.
- [39] Mosteller F., "Remarks on method of paired comparison: I. The least squares solution assuming equal standard deviation and equal correlation," *Psychometrika*, 16, 1951, 3-9.