SOUND QUALITY EVALUATIONS BASED ON HUMAN AUDITORY-BRAIN SYSTEM

YOSHIHARU SOETA [1] AND RYOTA SHIMOKURA [2]

[1] National Institute of Advanced Industrial Science and Technology (AIST). Health Research Institute, Osaka, Japan.
y.soeta@aist.go.jp
[2] Nara Medical University. Department of Otorhinolaryngology, Nara, Japan.
rshimo@naramed-u.ac.jp

Abstract – *A model of primary sensations and spatial sensations to environmental noise is proposed by Ando (2001). The model of the auditory–brain system includes the autocorrelation function (ACF) and the interaural cross-correlation function (IACF) mechanisms. At present, environmental noises are evaluated by sound intensity such as equivalent continuous A-weighted sound pressure level (L_{Aeq}). However, we sometimes feel annoyed with sound with low sound intensity because of the quality. Sound quality can be characterized by factors obtained from ACF and IACF of sounds. For example, pitch and pitch strength can be characterized by the delay time and amplitude of the maximum peak of the ACF. Directional sensation can be characterized by the delay time and amplitude of the maximum peak of the IACF. To verify the model, we investigated how ACF and IACF factors are coded in our human brain. The results indicated that the delay time and amplitude of the maximum peak of the ACF and IACF are coded by the latency, strength, and location of brain activity. In addition, we applied the model to evaluations of noises inside a train car. The results indicated that the effects of outside environments on noises inside the train car can be characterized by the ACF and IACF factors.*

1 INTRODUCTION

The main function of sensory systems is to get information about the outside world to the brain; thus it is necessary to consider how the information is processed in the human brain to evaluate the outer world. In hearing, that information is carried by pressure variations in the air (sound wave). At first, the sound wave pass through the peripheral auditory sysytem, the outer ear, middle ear, and inner ear.

The outer ear is the external part of the auditory system, including the pinnae and the ear canal. The pinnae significantly modifies incoming sound, particularly at high frequencies, and this is important in our ability to localize sounds. Sound travels down the ear canal and causes the eardrum, or tympanic membrane, to vibrate. Because of the resonance of the outer ear, we are more sensitive to sound frequencies between 1000 and 6000 Hz [1,2].

The middle ear is the air-filled space between the eardrum and the cochlea that contains the ossicles. The acoustic vaibrations on the eardrum are transmitted through the middle ear by three small bones, malleus, incus, and stapes, to the oval window of the cochlea. The middle ear acts as an impeadance-matching device or transformer that improves sound transmission and reduces the amount of reflected sound. This is accomplished mainly by the differences in effective areas of eardrum and the oval window, and to a small extent by the lever action of the ossicles. Transmission of sound through the middle ear is most efficient at frequencies between 500 and 4000 Hz [3,4].

The inner ear is the part of the ear that is filled with fluid, including the cochlea and semicircular canals. Sound enters the cochlea through the oval window covered by a membrane. When the oval window moves due to the pressure from the stapes, Reissner's membrane and the basilar membrane are pushed down, and the round window moves out. It follows that vibration of the stapes leads to vibration of the basilar membrane. The basilar membrane separates out the frequency components of a sound. At the base of the cochlea, near the oval window, the basilar membrane is narrow and stiff and is sensitive to high frequencies. The other end of the membrane, at the apex of the cochlea, is wide and loose and is sensitive to low freequencies. The basilar membrane behaves as a band of overlapping bandpass filters, which is called auditory filters.

The mechanical vibrations of the basilar membrane are converted into electrical activity in the auditory nerve. This task is accomplished by the inner hair cells. Vibration of the basilar membrane causes a displacement of a stereochilia at the tips of the hair cells which lies within the organ of Corti on the basilar membrane, and this leads to action potentials (spikes) within the nerve fibers of the auditory nerve. Because each inner hair cell is attached to a specific place on the basilar membrane, each neuron in the auditory nerve carries information about the vibration of the basilar membrane at a

single place in the cochlea. This means each neuron in the auditory nerve is sensitive to each characteristic frequency.

The auditory nerve carries the information about incoming sound from the cochlea to the cochlea nucleus. The information is passed via synapses to a number of other brainstem nuclei: the superior olivary complex, the lateral lemniscus, and inferior colliculus. Nerve fibers from the inferior colliculus synapse with the medial geniculate body, which is part of the thallamus in the midbrain. The thalamus acts as a sort of relay station for sensory information. Nerve fibers from the medical geniculate body project to the auditory cortex, which is part of the cerebral cortex and is invloved in basic sensory functions.

Environmental noises have been evaluated by sound intensity, such as equivalent continuous A-weighted sound pressure level (L_{Aeq}). However, we feel annoyed with sound with low intensity because of the quality [5]. To evaluate sound qualities, it is necessary to consider our auditory functionings, that is how incoming sound is processed from the peripheral to central auditory system. A model for evaluations of environmental noise has been proposed based on human auditory system [6]. The model of the auditory–brain system includes the autocorrelation function (ACF) mechanism, which might exist in the auditory nerve, and the interaural cross-correlation function (IACF) mechanism, which might exist in the inferior colliculus.

The parameteres obtained form ACF and IACF are useful for evaluation of sound quality. For example, pitch and pitch strength can be predicted by the delay time and amplitude of the maximum peak of the ACF [6,7]. Directional sensation can be predicted by the delay time and amplitude of the maximum peak of the IACF [8,9]. First, we report how ACF and IACF factors are coded in our human brain. Second, we show one application for the model based on ACF and IACF mechanisms in sound quality evaluation for noises inside a train car.

2 BRAIN ACTIVITIES RELATED TO PARAMETERS OF ACF AND IACF

2.1 Introduction

Pitch plays an important role in all aspects of hearing. Pitch phenomena can be explained in terms of the frequency composition or the time structure of the sound. In temporal models of pitch perception, it is assumed that the pitch is extracted with ACF [10-12]. Regarding the pitch and strength of pitch, psychophysical research has revealed that the pitch and the strength of the pitch corresponds well to the delay time and amplitude of the maximum peak of the ACF, τ_1 and ϕ_1, of the sound [7,13].

The most important cues for sound localization in human are the differences in intensity (ILD) and timing (ITD) of the sound waves received at the two ears. ITDs can be measured as τ_{IACC} by the IACF between two sound signals received at both the left and right ears. The width of the sound image changes according to the interaural cross-correlation (IACC) [14,15]. When sounds are delivered dichotically, the sound image varies with the IACC of the sound. If the IACC is high, the sound image is fused and occupies a narrow region. As the IACC decreases, the sound becomes more diffuse. Lateralization performance has been previously measured as a function of the degree of IACC [8, 9], and the results showed that lateralization performance decreases as the IACC is reduced.

We investigated how features of sound related to pitch, τ_1, pitch strength, ϕ_1, and, lateralization performance, IACC and τ_{IACC}, are represented in the human brain by Magnetoencephalography (MEG) (Fig. 1). MEG is a noninvasive technique for investigating neuronal activity in the living human brain. MEG measures the weak magnetic fields produced by electric currents flowing in neurons. The measured signals are generated by the synchronized neuronal activity related to sensory inputs, such as sound and light.

Figure 1: 122 channel MEG measurement system.

2.2 Methods

Ten normal-hearing participants participated in the experiment. Informed consent was obtained from each subject after the nature of the study was explained. The study was approved by the ethics committee of the National Institute of Advanced Industrial Science and Technology (AIST).

To control τ_1 and ϕ_1, the iterated rippled noise (IRN) was produced by a delay-and-add algorithm applied to the bandpass noise that was filtered using fourth-order Butterworth filters between 100-3500 Hz. The delay (τ_1) was set to 1, 2, 4, 8, 16, 32, and 64 ms, corresponding to frequency of 1000, 500, 250, 125, 62.5, 31.3, and 15.6 Hz, respectively. The number of iterations of the delay-and-add process was set to 2 and 32. The ϕ_1 increases as the number

of iterations increases. Examples of temporal waveforms, spectra, and ACF of some of the stimuli are shown in Fig. 2.

The IACC of the stimuli was controlled by mixing in-phase diotic bandpass and dichotic independent bandpass noises in appropriate ratios [16]. For stimulus lateralization, two cues were available to participants: envelope τ_{IACC} and ongoing τ_{IACC}. In this experiment, the envelope τ_{IACC} was zero for all stimuli, and the ongoing τ_{IACC} was varied. Here, "envelope" refers to the shape of a gating function with 10-ms linear ramps at the onset and offset. Examples of temporal waveforms, spectra, and IACFs of some of the stimuli are shown in Fig. 3.

The stimulus duration used the experiments was 0.5 s, including rise and fall ramps of 10 ms. Sound were presented binaurally to the left and right ears through an insert earphones with 29-cm plastic tubes and eartips inserted into the ear canals at SPL of 60 dB, and the ILD was set to 0 dB.

Brain activities evoked by the sound stimuli were recorded using a 122 channel whole-head magnetometer (Neuromag-122TM) in a magnetically shielded room. In order to maintain a constant vigilance level, the participants were instructed to concentrate on a self-selected silent movie that was being projected on a screen in front of them and to ignore the stimuli. Magnetic data were sampled at 400 Hz after being bandpass-filtered between 0.03 and 100 Hz, and then averaged approximately 100 times. The averaged responses were digitally filtered between 1.0 and 30.0 Hz. The analysis time was 0.7 s from 0.2 s prior to onset of the stimulus. The prestimulus period (average of the 0.2 s) was used as the baseline level. The peak latency having the maximum value of the root mean square (RMS) amplitudes in the latency range of 70-130 ms over each left and right hemisphere was defined as the latency of the N1m in each subject.

Source analysis based on the model of a single moving equivalent current dipole (ECD) in a spherical volume conductor was applied to the measured field distribution [17]. This gave estimates of the location and strength of auditory cortical activity. Source estimations were performed with a subset of 18-28 channels over each hemisphere around the latency of the N1m. The ECD location and moment were determined as the best fit that maximally accounts for the measured magnetic field distribution. These estimates were accepted for further analyses if the goodness of fit of the field of the estimated ECD to the measured magnetic field was >80%. We define the ECD moment as the amplitude of the N1m hereafter.

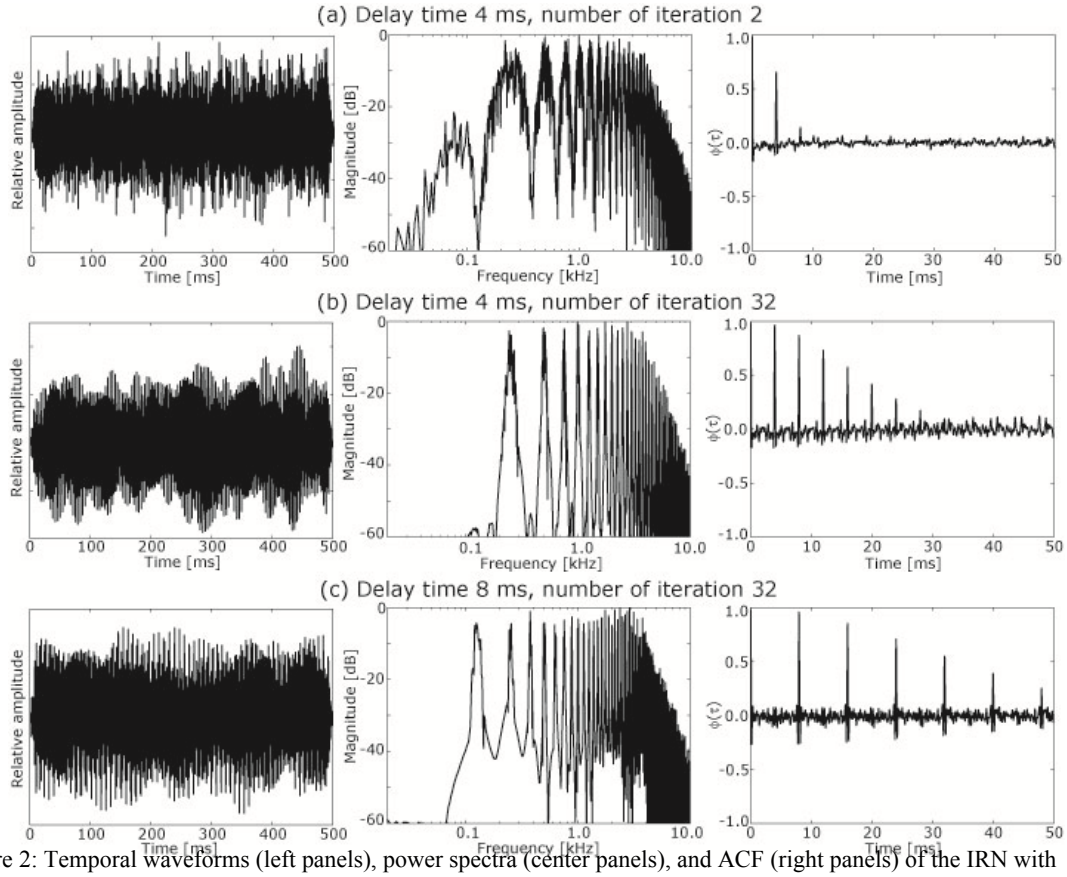

Figure 2: Temporal waveforms (left panels), power spectra (center panels), and ACF (right panels) of the IRN with different delay times (d, τ_1) and number of iterations (n). (a) d = 4 ms, n = 2; (b) d = 4 ms, n = 32; (c) d = 8 ms, n = 32.

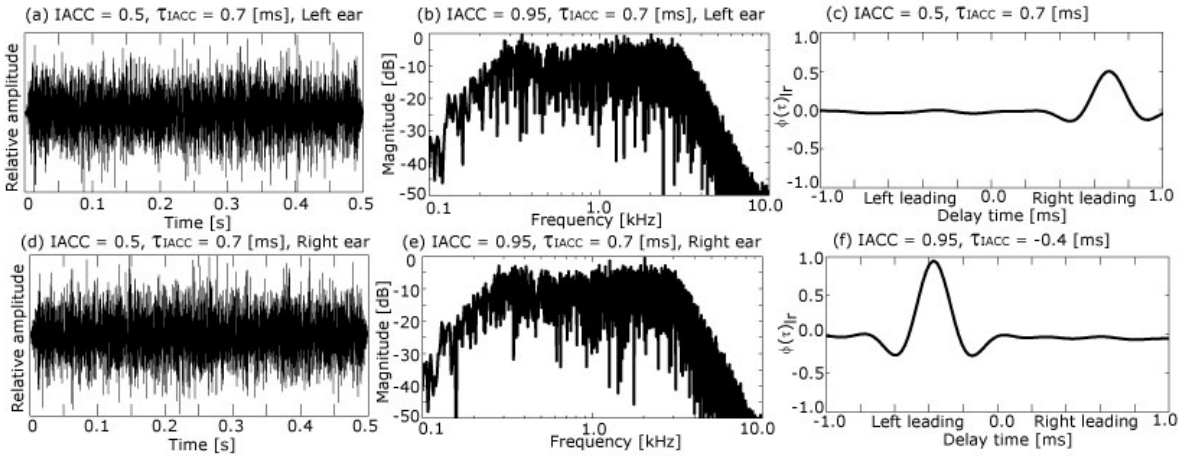

Figure 3: Temporal waveforms (left panels), power spectra (center panels), and IACF (right panels) of the bandpass noises with different IACC and τ_{IACC}.

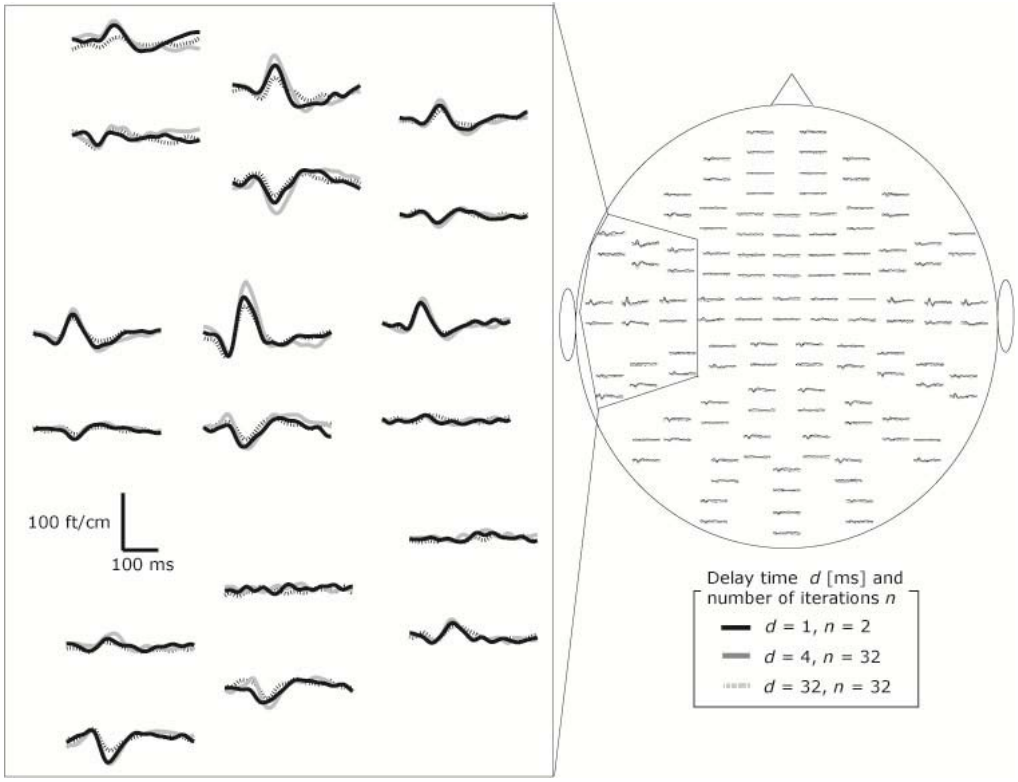

Figure 4: Typical waveforms of MEGs in response to IRNs from 122 channels in one participant.

2.2 Results and Discussion

All the stimuli elicited prominent N1m responses, which is found approximaely at 100 ms after the stimulus onset in both the left and right hemispheres, with the near-dipolar field patterns, as shown in Fig. 4. This indicaes the sources of the response was in the vicinity of the auditory cortex of each hemisphere.

Figure 5 shows the mean amplitude and latency of the N1m as a function of τ_1 (IRN delay) when the number of iteration was 2 and 32. The maximum amplitude of the N1m was produced by the 1-ms of τ_1, which correspond to a frequency of 1000 Hz. In previous studies, the N1m amplitude reached maximal levels at a frequency of approximately 1000 Hz and decreased at higher or lower frequencies for pure tone [18] and narrowband noises [19]. These results are consistent with the present findings. Moreover, for IRNs with 32 iterations, the N1m amplitude decreased sharply for τ_1 between 16 and 32 ms (diamonds in Fig. 5a). This is remarkably similar to the results of pitch onset response [20]. This suggests that the amplitude of the pitch-related response reflects the lower limit of the audible pitch

range [e.g., 21]. For IRNs with 2 iterations, the sharp decrease in the N1m amplitude was shifted towards smaller delays, which further corroborates the notion that the amplitude of the pitch-related response reflects pitch strength as a function of pitch. This conclusion is consistent with previous MEG [22-24] and functional magnetic resonance imaging (fMRI) results [25-27].

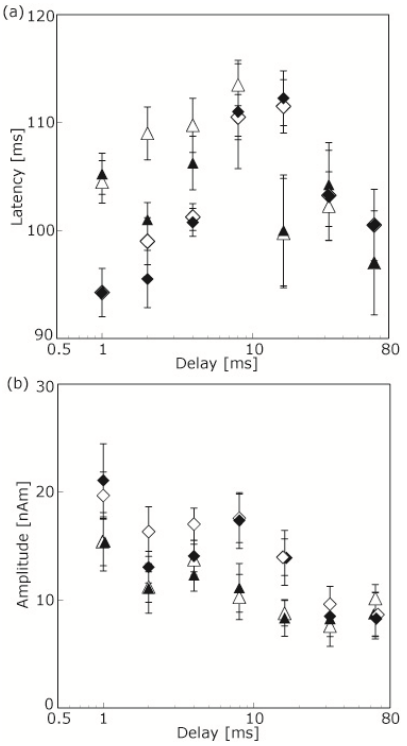

Figure 5: Mean (a) latency and (b) amplitude of the N1m response (± SEMs) as a function of τ_1 of the IRN. The number of iteration and hemisphere are indicated by △: 2, left; ▲: 2, right; ◇: 32, left; ◆: 32, right.

The longest latency of the N1m was observed at a τ_1 between 8 and 16 ms, whereas shorter and longer delays yielded relatively shorter N1m latencies (Fig. 5b). The current results are consistent with previous findings that the latency of the response to pure tones [22, 28-29], harmonic complex tones [22,30-32], and bandpass filtered noises [19] also tend to increase with decreasing τ_1 (pitch) up to some point.

Figure 6 shows the average ECD locations of ten participants for the different stimulus conditions. The ECD location of the N1m for τ_1 of 32 and 64 ms was slightly posterior compared to those for shorter delays. MEG responses evoked by regular and irregular click trains revealed that the sustained responses can be explained by the combination of an anterior source in lateral Heschl's gyrus, which is particularly sensitive to regularity and largely insensitive to intensity, and a posterior source in planum temporale, which is particularly sensitive to intensity and largely insensitive to regularity [33]. This is in good agreement with the present findings that the source of the N1m for τ_1 of 32 and 64 ms was slightly posterior compared to those for shorter delays. This suggests that at least part of the N1m generators are related to temporal pitch extraction that is based on τ_1, supporting the previous findings [32,34-36].

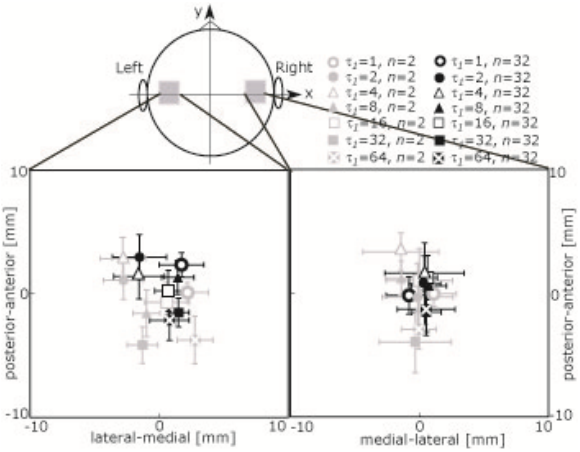

Figure 6: Mean ECD location (± SEM) on axial plane in response to the IRN with different τ_1 and numbers of iteration (*n*) in both left and right hemispheres.

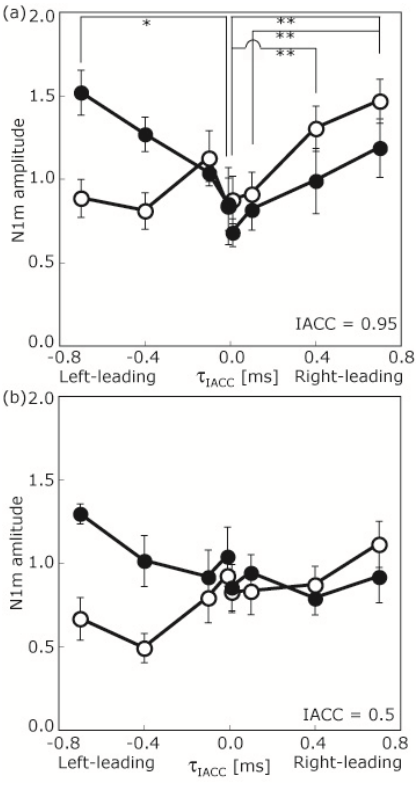

Figure 7: Mean N1m amplitude (± SEMs) as a function of the τ_{IACC} from the right (●) and left (○) hemispheres for (a) IACC = 0.95 and (b) IACC = 0.5. Asterisks indicate statistical significance (*P<0.05, **P<0.01; Post hoc Newman-Keuls test).

Figure 7 shows the N1m amplitude as a function of τ_{IACC}. When the IACC of the stimulus was 0.95, the N1m amplitude increased with increasing τ_{IACC} in the right hemisphere in the case of a left-leading stimulus, and in both the left and right hemispheres in the case of a right-leading stimulus. When the IACC of the stimulus was 0.5, the N1m amplitudes increased with increasing τ_{IACC} only in the right hemisphere in the case of a left-leading stimulus. This result is consistent with previous findings [36-38]. The N1m amplitude increased slightly with increasing τ_{IACCs} in the hemisphere contralateral to the τ_{IACCs} when the IACC of the stimulus was 0.5. Lateralization performance worsens with decreasing IACCs [8,9]; therefore, the present results may indicate that lateralization performance is reflected in N1m amplitudes. Put another way, there is a close relationship between the N1m amplitudes and the IACCs and τ_{IACCs} of the stimuli.

Figure 8 shows the averaged ECD locations in the left and right hemispheres. The location of the ECDs underlying the N1m responses did not vary as a function of τ_{IACC} or IACC, a finding in agreement with previous MEG results [36-37,39].

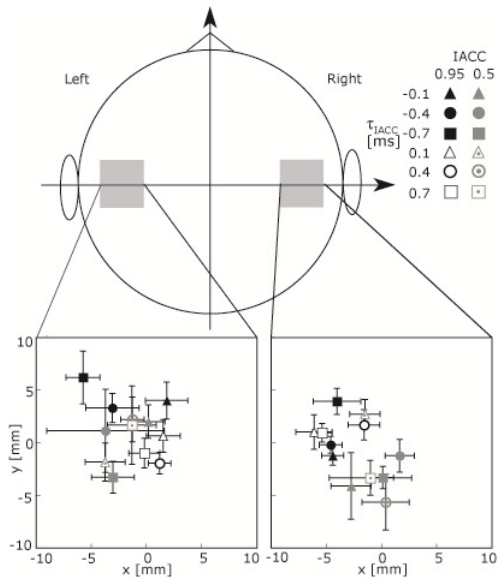

Figure 8: Mean ECD location (± SEM) of all subjects in the left and right temporal planes when the IACC was 0.95 and 0.5.

Recently it has been suggested that τ_{IACCs} may be coded by the activity level in two broadly tuned hemispheric channels [40-43]. The present study showed that the N1m amplitude varies with the τ_{IACC}; however, the location of the ECDs underlying the N1m responses did not vary with the τ_{IACC}. This suggests that different τ_{IACCs} are coded non-topographically but by response level. Thus, the current data seem to be more consistent with a two-channel model [40-43] rather than a topographic representation model [e.g., 44].

3 SOUND QUALITY ANALYSIS INSIDE A TRAIN CAR

3.1 Introduction

Trains are a major means of commuting and the average commute time is more than one hour in a typical metropolitan area of Japan. Consequently, working people and students spend long periods of time inside a train car. Noise in train cars can stress and fatigue passengers and interfere with their ability to understand the public address (PA) system. Therefore, the acoustic comfort of public transport systems should be maximized. In previous studies, only the sound pressure levels, such as L_{Aeq}, in trains have been measured [45]. However, measurements of the sound pressure level alone are not enough to characterize the acoustic comfort because other factors related to sound quality such as pitch, pitch strength, and diffuseness also affect the level of stress and fatigue passengers experience [5,6,46-48].

Although a great deal of research has been done on the noise emitted by the trains and its effect on nearby residents, there are relatively few studies on the noise in train cars. To improve the acoustic comfort and the ability of passengers to understand PA systems, it is necessary to clarify the characteristics of noises from a quantitative and qualitative point of view. In this study, octave-band power levels, L_{Aeq}, and IACF/ACF parameters were used to characterize the noises in train cars.

3.2 Methods

Noise was measured continually in a train car running from one terminal station to another terminal station on three railway lines. Some intervals between stations were selected for analysis according to the outer environment of the train car, that is, above ground and in four type of tunnels. The tunnels of the underground trains are distinguished according to the excavation methods. The cut-and-cover method (CCM) excavates the tunnel from ground level downwards, constructs the tunnel roof and sides and buries it. The cross-section of the tunnel is rectangular and large. The boring machine method (BMM) drills tunnels using a boring machine and lines the insides with steel segments. The cross-section of the tunnel is circular and small. The new Austrian tunneling method (NATM) excavates a tunnel that is maintained by the retentive characteristics of the natural ground, the surface of which is strengthened with shotcrete and the insertion of reinforcing lock bolts. The cross-sectional area of the tunnel lies between CCM and BMM.

The dummy head microphone (KU100, Nuemann) were used to record noises inside a train

car. The dummy head is an artificial model of the human head with two microphones inserted in the location of the eardrums. It can simulate the frequency-dependent distortions of phase and amplitude on sound reaching the entrance of the left and right ear canals. The microphones were located in the area for a wheelchair user. The microphones were always facing inside the train cars and perpendicular to the direction of travel. The heights of the dummy head microphones were 1.6 m. For all measurements, noise was recorded on a personal computer (Let's note, Panasonic) via an AD/DA converter (AudioFire8, Echo Digital Audio) at a sampling rate of 48 kHz and a sampling resolution of 32 bits. The measurements were carried out in the daytime (10:00 to 15:00 hours) to avoid the rush hours.

The octave-band power levels, L_{Aeq}, and parameters extracted from the IACF/ACF were derived from the left and right output signals $p_l(t)$ and $p_r(t)$ from the dummy head microphone. The L_{Aeq} is widely used for the measurement of noise [49]. The octave-band power levels and L_{Aeq} were determined from the time-averaged sound pressure levels of the octave-band filtered and A-weighted $p_l(t)$ and $p_r(t)$.

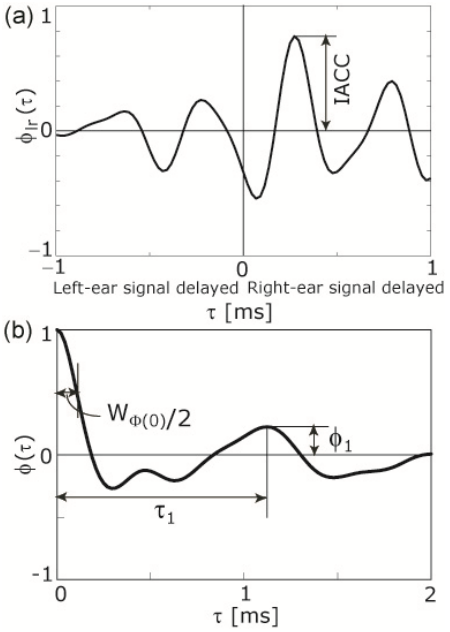

Figure 9: Definitions of parameters extracted from (a) IACF and (b) ACF.

The IACF/ACF parameters of noise are proposed for describing the sound quality [6]. The definitions of IACF/ACF parameters are shown in Fig. 9. The normalized IACF between the signals received at each ear, $p_l(t)$ and $p_r(t)$, as a function of running step, s, is defined by:

$$\phi_{lr}(\tau) = \phi_{lr}(\tau,s,T) = \frac{\Phi_{lr}(\tau;s,T)}{\sqrt{\Phi_{ll}(0;s,T)\Phi_{rr}(0;s+\tau,T)}}, \quad (1)$$

where:

$$\Phi_{lr}(\tau;s,T) = \frac{1}{2T}\int_{s-T}^{s+T} p'_l(t)\,p'_r(t+\tau)dt, \quad (2)$$

where $\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ are the ACFs at $\tau = 0$ for the left and right ear, respectively, 2T is the integration interval, and $p'_l(t)$ and $p'_r(t)$ are obtained from $p_l(t)$ and $p_r(t)$ after passing through the A-weighted network. The interaural cross-correlation coefficient (IACC) is defined as the maximum correlation of the sounds arriving at the left and right ears, and is related to the subjective diffuseness [15]. In a similar way, the normalized ACF of a signal received at an ear, $p_l(t)$ and $p_r(t)$, can be obtained by substituting $p'_l(t)$ ($p'_r(t)$) for $p'_r(t)$ ($p'_l(t)$) in Eq (2). The ACF parameters, τ_1 and ϕ_1, are defined as the delay time and the amplitude of the first maximum peak and are related to the perceived pitch and pitch strength [6,7]. The other ACF parameter, $W_{\Phi(0)}$, is defined as the width of the first decay and corresponds to the spectral centroid [50].

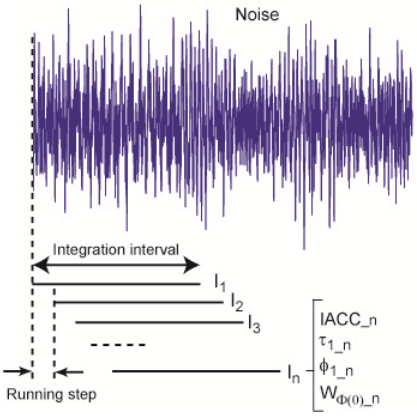

Figure 10: Concept of short-time moving ACF analysis along the noise.

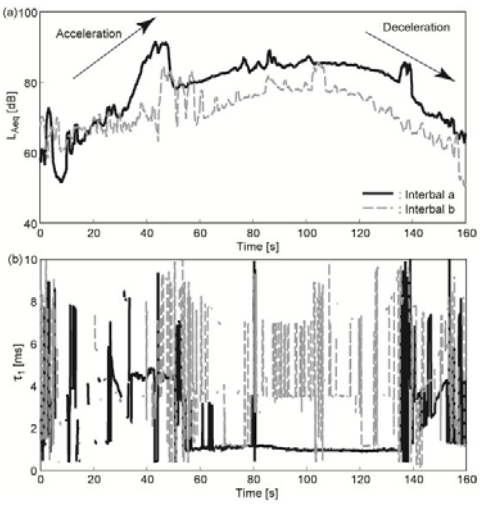

Figure 11: Examples of factors (a) L_{Aeq} and (b) τ_1 as a function of time in an interval between stations.

The concept of a short-time moving analysis is illustrated in Fig. 10. The IACC, τ_1, ϕ_1, and $W_{\Phi(0}$ were calculated at every given integration interval 2T. The start of each analysis was delayed for a short time, which is the running step. The moving analysis aids in the effective description of the temporal change of the parameters. In this study, the integration interval, 2T, was 0.5 s, and the running step was 0.1 s. Figure 11 shows examples of parameters L_{Aeq} and τ_1 as a function of time. These temporal variations are very complicated, because there are so many kinds of noises inside a train car, such as rolling, curve squeal, impact noises. Then, we do not focus on each noise afterward.

3.3 Results and Discussion

Figure 12 shows the L_{Aeq} and octave-band power levels. The L_{Aeq} in the tunnels was 5-7 dB higher than that above ground. The reflections in the tunnels clearly increased the L_{Aeq}. In all types of tunnel, the power level had a peak around 250 Hz. IN contrast, the power level had a dip around 250 Hz above the ground. The difference between the power levels in tunnels and above ground was large at frequency bands around 250 and 500 Hz. These results suggest the effect of reflections in tunnels can be mainly observed around 250 and 500 Hz.

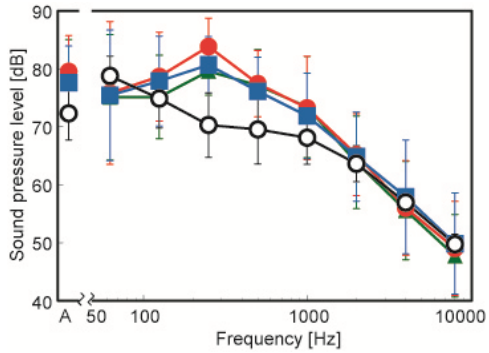

Figure 12: Measured sound pressure level as a function of the 1/1 octave band centre frequency and L_{Aeq} in (■) CCM, (●) BMM, (▲) NATM tunnels, and (○) above ground.

On the basis of a running analysis, the cumulative frequencies of IACC, τ_1, ϕ_1, and $W_{\Phi(0)}$ in each outside environment were obtained (Fig. 13). The IACC was largest in the BMM tunnel, smaller in NATM tunnels, and smallest above ground. More reflections were evident inside train cars in the BMM tunnels because of the smaller circular cross-section. In contrast, fewer reflections were evident inside train cars in the NATM tunnels because of the complex cross-section. The values of τ_1 were concentrated at 1 ms, which corresponds to a frequency of 1000 Hz, in the NATM tunnel and above ground. The noise at 1000 Hz could be rolling noise emitted from the wheels and rails [51,52]. The values of τ_1 were concentrated at 4 ms, which corresponds to a frequency of 250 Hz, in the CCM, BMM, and NATM tunnels. The ϕ_1 value was largest in the NATM tunnel, smaller in BMM and CCM tunnels, and smallest above ground. The reflections in tunnels could cause the increase of ϕ_1 values and the perceived pitch strength of noises inside the train cars might be stronger in the tunnels and cause more annoyance [53,54]. The $W_{\Phi(0)}$ was largest in the BMM tunnel, smaller in CCM and NATM tunnels, and smallest above ground. This suggests that the tunnel makes the spectral centroid of noises inside the train cars lower, especially in the BMM tunnels.

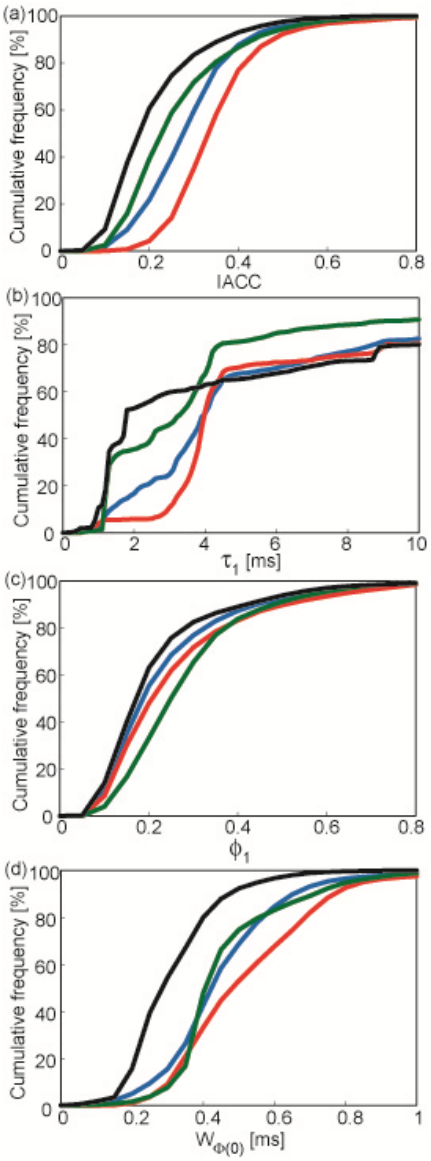

Figure 13: Cumulative frequencies of (a) IACC, (b) τ_1, (c) ϕ_1, and (d) $W_{\Phi(0)}$ obtained from short-time moving analyses in (—) CCM, (—) BMM, (—) NATM tunnels, and (—) above ground.

4 CONCLUSIONS

First, we investigated how ACF and IACF factors are coded in our human brain by MEG. ACF factors, τ_1 and ϕ_1, are reflected in the latency, amplitude, and source location of the evoked response, i.e., N1m. IACF factors, IACC and τ_{IACC}, are reflected in the amplitude of the evoked response. This suggests human perception is based on these parameters. Second, we applied the model based on ACF and IACF mechanisms to sound quality evaluation for noises inside a train car. ACF and IACF parameters, such as τ_1, ϕ_1, $W_{\Phi(0)}$, and IACC can quantify dominant frequency components, strength of the frequency components, and spectral centroid of the noises, degree of diffusion inside a train car. Therefore, parameters extracted from the ACF and IACF could be useful criteria for evaluating quality of environmental noise.

5 REFERENCES

[1] Shaw, E.A.G. "*Transformation of sound pressure level from the free field to the eardrum in the horizontal plane*". J. Acoust. Soc. Am. Vol. 56, pp. 1848-1861. 1974.

[2] Mehrgardt, S., Mellert, V. "*Transformation characteristics of the external human ear*". J. Acoust. Soc. Am. Vol. 61, pp. 1567-1576. 1977.

[3] Puria, S., William, T., Peake, W.T., Rosowski, J.J. "*Sound-pressure measurements in the cochlear vestibule of human-cadaver ears*". J. Acoust. Soc. Am. Vol. 101, pp. 2754-2770. 1997.

[4] Aibara, R., Welsh, J.T., Puria, S., Goode, R.L. "*Human middle-ear sound transfer function and cochlear input impedance*". Hear. Res. Vol. 152, pp. 100-109. 2001.

[5] Kitamura, T., Shimokura, R., Sato, S., Ando, Y. "*Measurement of Temporal and Spatial Factors of a Flushing Toilet Noise in a Downstairs Bedroom*". J. Temporal Des. Arch. Environ. Vol. 2, pp 13-19. 2002.

[6] Ando, Y. "*A theory of primary sensations and spatial sensations measuring environmental noise*". J. Sound Vib. Vol. 241, pp. 3-18. 2001.

[7] Yost, W.A. "*Pitch strength of iterated ripple noise*". J. Acoust. Soc. Am., Vol. 100, pp. 3329-3335. 1996.

[8] Jeffress, L.A., Blodgett, H.C., Deatherage, B.H. "*Effects of interaural correlation on the precision of centering a noise*". J. Acoust. Soc. Am. Vol. 34, pp. 1122-1123. 1962.

[9] Zimmer, U., Macaluso, E. "*High binaural coherence determines successful sound localization and increased activity in posterior auditory areas*". Neuron, Vol. 47, pp. 893-905. 2005.

[10] Licklider, J.C.R. "*A duplex theory of pitch perception*". Experimenta, Vol. 7, pp. 128-134. 1951.

[11] Meddis, R., Hewitt, M.J. "*Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification*". J. Acoust. Soc. Am. Vol. 89, pp. 2866-2882. 1991.

[12] Cariani, P.A., Delgutte, B. "*Neural correlates of the pitch of complex tones. I. Pitch and pitch salience*". J. Neurophysiol. Vol. 76, pp. 1698-1716. 1996,

[13] Wightman, F.L. "*The pattern-transformation model of pitch*". J. Acoust. Soc. Am. Vol. 54, pp. 407-416. 1973.

[14] Licklider, J.C.R. "*The influence of interaural phase relations upon masking of speech by white noise*". J. Acoust. Soc. Am. Vol. 20, pp. 150-159. 1948.

[15] Ando, Y., Kurihara, Y. "*Nonlinear response in evaluating the subjective diffuseness of sound fields*". J. Acoust. Soc. Am. Vol. 80, pp. 833-836. 1986.

[16] Blauert, J., Lindemann, W. "*Spatial mapping of intracranical auditory events for various degrees of interaural coherence*". J. Acoust. Soc. Am. Vol. 79, pp. 806-813. 1986.

[17] Hämäläinen, M.S., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V. "*Magnetoencephalography – theory instrumentation and applications to noninvasive studies of the working human brain*". Rev. Mod. Phys. Vol. 65, pp. 413–497. 1993.

[18] Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., Elbert, T. "*Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings*". Electroenceph. Clin. Neurophysiol. Vol. 94, pp. 26-40. 1995.

[19] Soeta, Y., Nakagawa, S., Matsuoka, K. "*The effect of center frequency and bandwidth on the auditory evoked magnetic field*". Hear. Res. Vol. 218, pp. 64-71. 2006.

[20] Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., Lütkenhöner, B. "*Neuromagnetic evidence for a pitch processing center in Heschl's gyrus*". Cereb. Cortex Vol. 13, pp. 765-772. 2003.

[21] Pressnitzer, D,, Patterson, R.D., Krumbholz, K. "*The lower limit of melodic pitch*". J. Acoust. Soc. Am. Vol. 109, pp. 2074-2084. 2001.

[22] Seither-Preisler, A., Krumbholz, K., Lutkenhoner, B. "Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: A function of the auditory periphery?". Audiol. Neurootol. Vol. 8, pp. 322-337. 2003.

[23] Soeta, Y., Nakagawa, S., Tonoike, M. "*Auditory evoked magnetic fields in relation to bandwidth variations of bandpass noise*". Hear. Res. Vol. 202, pp. 47-54. 2005.

[24] Soeta, Y., Nakagawa, S., Tonoike, M. "*Auditory evoked magnetic fields in relation to the iterated rippled noise*". Hear. Res. Vol. 205, pp. 256-261. 2005.

[25] Griffiths, T.D., Buechel, C., Frackowiak, R.S.J., Patterson, R.D. "*Analysis of temporal structure in*

sound by the human brain". Nat. Neurosci. Vol. 1, pp. 421-427. 1998.
[26] Griffiths, T.D., Uppenkamp, S., Johnsrude, I., Josephs, O., Patterson, R.D. "*Encoding of the temporal regularity of sound in the human brainstem*". Nat. Neurosci. Vol. 4, pp. 633-637. 2001.
[27] Penagos, H., Melcher, J.R., Oxenham, A.J. "*A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging*". J. Neurosci. Vol. 24, pp. 6810-6815. 2004.
[28] Roberts, T.P.L., Poeppel, D. "*Latency of auditory evoked M100 as a function of tone frequency*". Neuroreport Vol. 7, pp. 1138-1140. 1996.
[29] Stufflebeam, S.M., Poeppel, D., Rowley, H.A., Roberts, T.P.L. "*Peri-threshold encoding of stimulus frequency and intensity in the M100 latency*". Neuroreport Vol. 9, pp. 91-94. 1998.
[30] Crottaz-Herbette, S., Ragot, R. "*Perception of complex sounds: N1 latency codes pitch and topography codes spectra*". Clin. Neurophysiol. Vol. 111, pp. 1759-1766. 2000.
[31] Langner, G., Sams, M., Heil, P., Schulze, H. "*Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalograph*y". J. Comp. Physiol. A. Vol. 181, pp. 665-676. 1997.
[32] Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., Lutkenhoner, B. "*Evidence of pitch processing in the N100m component of the auditory evoked field*". Hear. Res. Vol. 213, pp. 88-98. 2006.
[33] Gutschalk, A., Patterson, R.D., Rupp, A., Uppenkamp, S., Scherg, M. "*Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex*". NeuroImage Vol. 15, pp. 207-216. 2002.
[34] Gutschalk, A,, Patterson, R.D., Scherg, M., Uppenkamp, S., Rupp, A. "*Temporal dynamics of pitch in human auditory cortex*". NeuroImage Vol. 22, pp. 755-766. 2004.
[35] Seither-Preisler, A., Krumbholz, K., Patterson, R.D., Seither, S., Lutkenhoner, B. "*Interaction between the neuromagnetic responses to sound energy onset and pitch onset suggests common generators*". Eur. J. Neurosci. Vol. 19, pp. 3073-3080. 2004.
[36] Sams, M., Hämäläinen, M., Hari, R., McEvoy, L. "*Human auditory cortical mechanisms of sound lateralization: I. Interaural time differences within sound*". Hear. Res. Vol. 67, pp. 89-97. 1993.
[37] McEvoy, L., Hari, R., Imada, T., Sams, M. "*Human auditory cortical mechanisms of sound lateralization: II. Interaural time differences at sound onset*". Hear. Res. Vol. 67, pp. 98-109. 1993.
[38] Palomäki, K., Tiitinen, H., Mäkinen, V., May, P. J. C., Alku, P. "*Spatial processing in human auditory cortex: The effects of 3D, ITD, and ILD stimulation techniques*". Cogn. Brain Res. Vol. 24, pp. 364-379. 2005.
[39] Soeta, Y., Hotehama, T., Nakagawa, S., Tonoike, M., Ando, Y. "*Auditory evoked magnetic fields in relation to the inter-aural cross-correlation of bandpass noise*". Hear. Res. Vol. 196, pp. 109-114. 2004.
[40] McAlpine, D., Jiang, D., Palmer, A.R. "*A neural code for low-frequency sound localization in mammals*". Nat. Neurosci. Vol. 4, pp. 396-401. 2001.
[41] Brand, A., Behrend, O., Marquardt, T., McAlpine, D., Grothe, B. "*Precise inhibition is essential for microsecond interaural time difference coding*". Nature Vol. 417, pp. 543-547. 2002.
[42] McAlpine, D., Grothe, B. "*Sound localization and delay lines – do mammals fit the model?*". Trends Neurosci. Vol. 13, pp. 347-350. 2003.
[43] Stecker, G.C., Harrington, I.A., Middlebrooks, J.C. "*Location coding by opponent neural populations in the auditory cortex*". PLoS Biol. Vol. 3, pp. 520-528. 2005.
[44] Jeffres, L.A. "*A place theory of sound localization*". J. Comp. Physiol. Psychol. Vol. 41, pp. 35-39. 1948.
[45] Kono, S., Sone, T., Nimura, T. "*Personal reaction to daily noise exposure*". Noise Control Eng. J. Vol. 19, pp. 4-16. 1982.
[46] Fujii, K., Atagi, J., Ando, Y. "*Temporal and spatial factors of traffic noise and its annoyance*". J. Temporal Des. Arch. Environ. Vol. 2, pp. 33-41. 2002.
[47] Sato, S., You J., Jeon, J.Y. "*Sound quality characteristics of refrigerator noise in real living environments with relation to psychoacoustical and autocorrelation function parameters*". J. Acoust. Soc. Am. Vol. 122, pp. 314-325. 2007.
[48] Jeon, J.Y., Sato, S. "*Annoyance caused by heavy-weight floor impact sounds in relation to the autocorrelation function and sound quality metrics*". J. Sound Vib. Vol. 311, pp. 767-785. 2008.
[49] ISO 3381, "*Railway applications – Acoustics-Measurement of noise inside railbound vehicle*". International Organization for Standardization, 2005.
[50] Ando, Y., Cariani, P. "*Auditory and Visual Sensations*". Springer, New York, 2009.
[51] Stüber, C. "*Air- and structure-borne noise on railways*". J. Sound Vib. Vol. 43, pp. 281–289. 1975.
[52] Thompson, D.J., Jones, C.J.C. "*A review of the modelling of wheel/rail noise generation*". J. Sound and Vib. Vol. 231, pp. 519-536. 2000.
[53] Sato, S., Kitamura, T., Ando, Y. "*Loudness of sharply 2068 dB/octave filtered noises in relation to the factors extracted from the autocorrelation function*". J. Sound Vib. Vol. 250, pp. 47–52. 2002.
[54] Soeta, Y., Maruo, T., Ando, Y. "*Annoyance of bandpass filtered noises in relation to the factor extracted from autocorrelation function*". J. Acoust. Soc. Am. Vol. 116, pp. 3275-3278. 2004.