Effect of the Repetitive Components of a Noise on Loudness

Yoshiharu Soeta and Seiji Nakagawa
Institute for Human Science and Biomedical Engineering
National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31 Midorigaoka, Ikeda, Osaka, 563-8577 Japan

(Received 31 January 2008; accepted 6 October 2008)

To determine the effect of the repetitive component of a noise, which is characterized by the effective duration of the autocorrelation function, τ_e, the loudness of iterated rippled noises (IRNs) with different numbers of iterations and delay times under conditions of equal sound pressure was examined. Loudness matches were obtained using a two-interval, adaptive forced-choice procedure converging on the point of subjective equality following a simple 1-up, 1-down rule. Results indicated that the loudness of IRNs increased with increasing τ_e, from 10 to 100 ms, indicating the repetitive components of noises could affect their loudness in certain ranges of τ_e. The increase in loudness for the τ_e values from 10 and 100 ms corresponded to an approximately 5 dB increase in SPL.

Key words: Loudness, Autocorrelation function, Critical band, Pitch, Pitch strength

1. INTRODUCTION

Changes in frequency, duration and bandwidth all affect the perceived loudness of a stimulus, even when the intensity is constant (Yost 2000). Previous studies have concluded that the loudness of a noise remains constant as its bandwidth increases up to the critical band; thereafter, the loudness increases with increasing bandwidth beyond the critical band (Zwicker et al. 1957; Greenwood 1961; Scharf 1962). However, the loudness of a sharply (2068 dB/octave) filtered bandpass noise (BN) increases as the effective duration of the autocorrelation function (ACF), τ_e, increases, even when the bandwidth of the BN is within the critical band (Sato et al. 2002; Soeta et al. 2004). The τ_e represents repetitive components within the signal itself and increases as the bandwidth of a BN decreases. However, the envelope and sound-pressure level (SPL) also vary as the bandwidth of a BN changes. This variation of the envelope and SPL might therefore affect the loudness of a BN (Zhang and Zeng 1997; Moore et al. 1999). To eliminate such effects, our previous study investigated the effects of τ_e on loudness using iterated rippled noise (IRN) (Soeta et al. 2007). IRN is produced by delaying a noise, adding it to the original, and iterating the delay-and-add process. The reciprocal of the delay determines the pitch, and the number of iterations determines the pitch strength (Yost 1996). Thus, the delay time and number of iterations determine the τ_e of the ACF. In the previous study, the scale values of the loudness of IRNs for each delay, i.e., 1, 2, 4 and 8 ms, were obtained using a paired-comparison method (Soeta et al. 2007). We reported that the averaged scale values of the loudness of the IRNs for each delay increased with increasing τ_e of the ACF. Because subjects did not compare the loudness of IRNs with different delays and because the scale values of the loudness are regarded as the linear psychological distance between the IRNs that have been compared for each delay, it is impossible to compare the scale values for IRNs with different delays and evaluate the increase in loudness in dB scale. In the present study, therefore, loudness matches were obtained using a two-interval, adaptive forced-choice procedure converging on the point of subjective equality (PSE) following a simple 1-up, 1-down rule.

2. METHOD

The ACF provides the same information as the power spectral density of a signal. A normalized ACF can be expressed by

$$\phi(\tau) = \frac{\Phi(\tau)}{\Phi(0)}, \tag{1}$$

where

$$\Phi(\tau) = \frac{1}{2T}\int_0^{2T} p(t)\,p(t+\tau)\,dt \,, \qquad (2)$$

in which 2T is the integral interval, τ is the time delay and $p(t)$ is the signal as a function of time. The following can be determined from the ACF analysis: (1) the energy represented at zero delay, $\Phi(0)$; (2) the τ_e, defined as the time delay at which the envelope of the normalized ACF becomes 0.1; (3) the amplitude of the first non-zero maximum peak at positive delay, ϕ_1; and (4) the delay time τ_1 (Ando 1998; Ando 2001). The $\Phi(0)$ corresponds to the SPL. The τ_1 of the IRN corresponds to the delay of the IRN. The τ_e and ϕ_1 increase as the number of iterations increases. Although the ϕ_1 did not change as a function of the delay time of the IRN; the τ_e did change.

Digitally generated white noise with a sampling rate of 48 kHz was used to produce BN. IRN was produced by a delay-and-add algorithm applied to the BN that was filtered using fourth-order Butterworth filters between 100-3500 Hz. The number of iterations of the delay-and-add process was set at 2, 4, 8, 16 and 32. The delay was set at 0.5, 1, 2, 4, 8 and 16 ms, for which the reciprocals were 2000, 1000, 500, 250, 125 and 62.5 Hz, respectively. The duration of the stimuli was 0.5 seconds, including linear rise and fall ramps of 10 ms. The sounds were digital-to-analogue converted with a 16-bit sound card and sampling rate of 48 kHz. The sounds were presented diotically at a sound pressure level (SPL) of 60 dB through insert earphones (Etymotic Research ER-2) with 29-cm plastic tubes and eartips inserted into the ear canals. The passband in the transfer function of the plastic tubes approximately corresponded to the passband of the stimuli (100-3500 Hz). Figure 1 shows the temporal waveforms, power spectra and ACFs of some of the stimuli. Figure 2 shows the τ_e and ϕ_1 (calculated at integration interval 0.5 s)

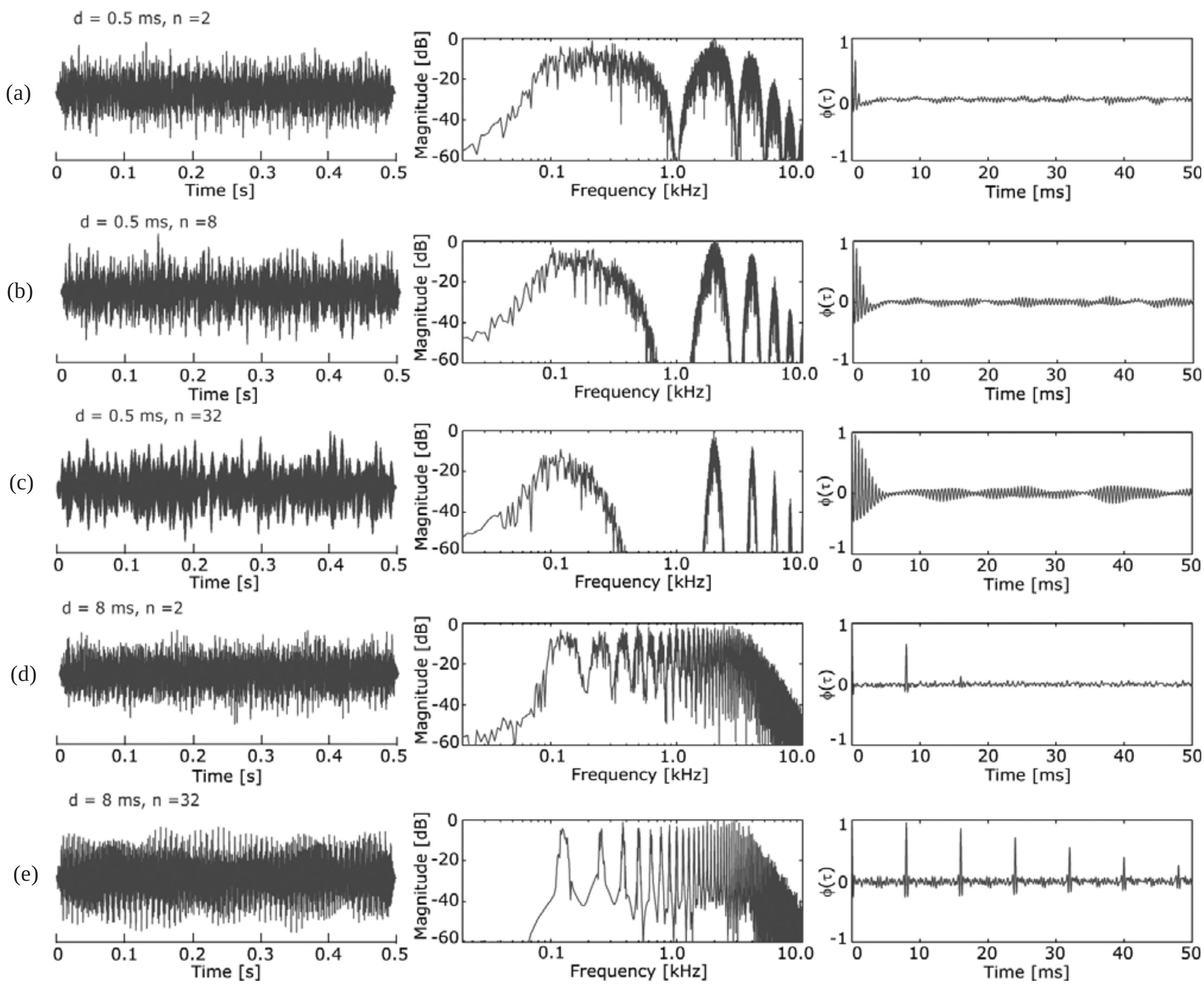

Fig. 1. Temporal waveforms (left panels), power spectra (middle panels) and ACF (right panels) of IRN with different delay times (d) and number of iterations (n). (a) d = 0.5 ms, n = 2; (b) d = 0.5 ms, n = 8; (c) d = 0.5 ms, n = 32; (d) d = 8 ms, n = 2; (e) d = 8 ms, n = 32.

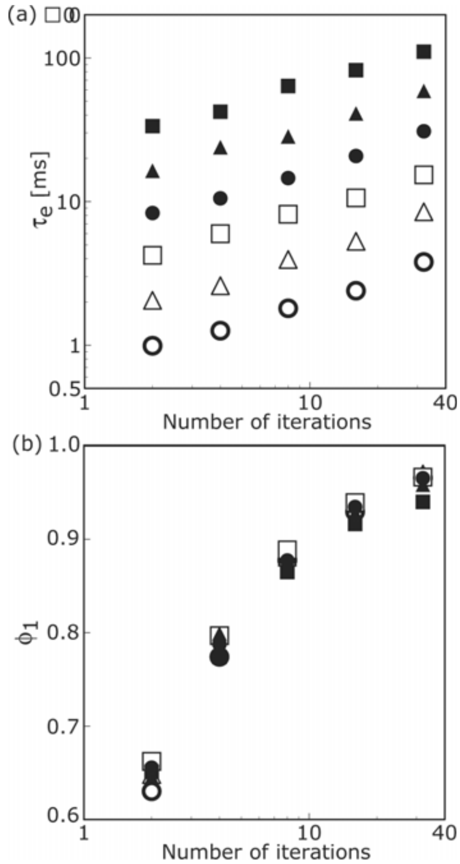

Fig. 2. (a) τ_e and (b) ϕ_1 of the stimuli used in the experiment as a function of the number of iterations with delays of (○) 0.5, (△) 1, (□) 2, (●) 4, (▲) 8 and (■) 16 ms.

of the stimuli used in the experiment. The integration interval was determined according to the psychological present, i.e., 0.5-5.0 s (Fraisse 1984).

Ten subjects (aged 21-37 years) with normal hearing took part in the experiment. The participants all had normal audiological status and no history of neurological diseases. There was no practice session because all subjects had already performed the same task to obtain loudness matches. Informed consent was obtained from each subject after the nature of the study had been explained. The protocol was approved by the ethics committee of the National Institute of Advanced Industrial Science and Technology (AIST), Japan.

Loudness matches were obtained using a two-interval, adaptive forced-choice procedure converging on the PSE following a simple 1-up, 1-down rule (Levitt 1971) in an anechoic and soundproof room. In each trial, the fixed (test) and variable (reference) sounds were presented in random order with equal, a priori probability, and separated by a 500-ms pause. The test sound was an IRN and the reference sound was a 1-kHz pure tone. Each subject's task was to indicate which sound was louder by pressing a key. For each adaptive track, the overall level of the test sound was fixed at 60 dB SPL, and the starting level of the reference sound was 50 dB SPL. The level of the reference sound was controlled by an adaptive procedure: whenever the subject judged the reference sound to be louder than the test sound, its SPL was lowered by a given amount; whenever the subject judged the test sound to be louder than the reference sound, the SPL of the reference sound was increased by that same amount. The initial step size was 5 dB; after four reversals (that is, changes in the direction of the adaptive track), it was decreased to 2 dB. A total of 12 reversals were collected for each adaptive track; the arithmetic mean of the last four was used to estimate the PSE. One estimate was obtained for each IRN and subject.

The effects of stimulus parameters (number of iterations and delay time) on the PSE for loudness were statistically analyzed by a repeated-measures analysis of variance (ANOVA) with two within-subject factors: number of iterations×delay time. Where appropriate, probabilities were adjusted by the Greenhouse-Geisser correction.

3. RESULTS

Figure 3 shows a typical example of a subject's response by simple 1-up 1-down procedures. This clearly shows how well the response converged to the PSE. The results were similar for all subjects, that is, all responses converged well to the PSE.

Figure 4 shows the PSE for loudness as a function of the number of iterations of the IRN for each delay time. The main effects of the number of iterations [$F(4, 36) = 7.94$, $P < 0.001$] and delay time [$F(5, 45) = 8.95$, $P < 0.005$, $\varepsilon = 0.36$] were significant. The interaction was also significant [$F(20, 180) = 5.73$, $P < 0.001$]. Hence, the main effects of the number of iterations were analyzed by a one-way repeated-

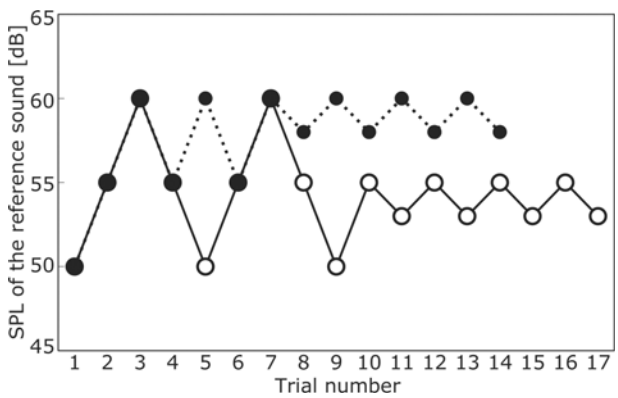

Fig. 3. Typical data for simple up-down procedures. A typical set of data from a subject in response to IRN with a delay time of 4 ms and (○) 8 and (●) 32 iterations.

measures ANOVA for each delay time.

When the delay was 0.5 ms, the averaged PSE for loudness decreased as the number of iterations increased. Seven of the 10 subjects had decreasing PSE for loudness with increasing number of iterations. The main effect of the number of iterations on the PSE for loudness was significant [$F(4, 36) = 5.19$, $P < 0.005$].

When the delay was 1 ms, the averaged PSE for loudness decreased as the number of iterations increased when the number of iterations was less than 8; however, the averaged PSE for loudness slightly increased with increasing number of iteration when the number of iterations was 16 and 32. Seven of the 10 subjects showed these tendencies. The main effect of the number of iterations on the PSE for loudness was significant [$F(4, 36) = 2.75$, $P < 0.05$].

When the delay was 2 ms, the averaged PSE for loudness increased as the number of iterations increased, except for when the number of iterations was 2. Seven of the 10 subjects showed such tendencies. The main effect of the number of iterations on the PSE for loudness was significant [$F(4, 36) = 5.64$, $P < 0.005$].

When the delay was 4 and 8 ms, the averaged PSE for loudness increased as the number of iterations increased, except for when the number of iterations was 32. Seven of the

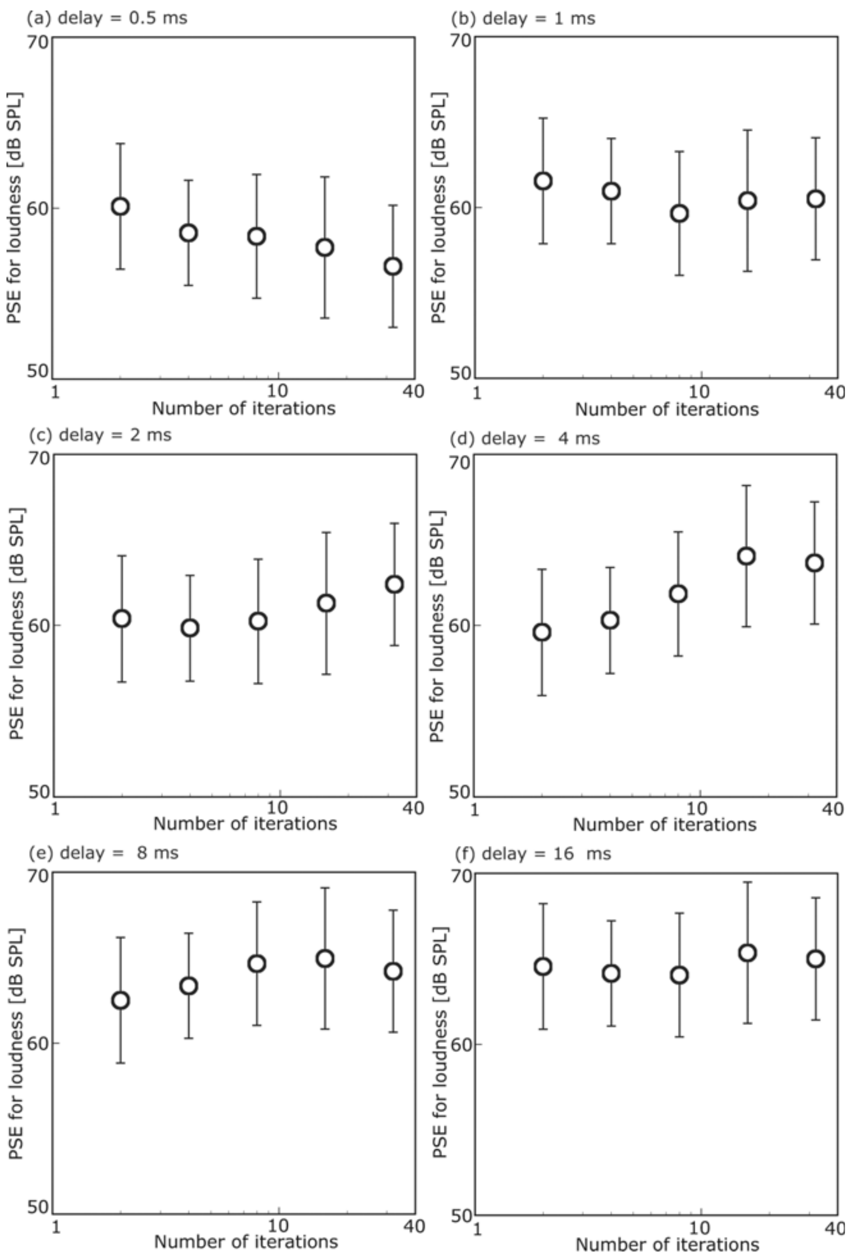

Fig.4. Mean PSE for loudness as a function of the iteration number of IRN with a delay of (a) 0.5, (b) 1, (c) 2, (d) 4, (e) 8 and (f) 16 ms. Error bars represent the standard deviation of the means.

10 subjects showed such tendencies. Three of the 10 subjects did not show clear PSE change as a function of number of iterations. The main effect of the number of iterations on the PSE for loudness was significant for the delay of 4 ms [$F(4, 36) = 11.27$, $P < 0.001$] and the delay of 8 ms [$F(4, 36) = 8.14$, $P < 0.001$].

When the delay was 16 ms, the averaged PSE for loudness did not change significantly [$F(4, 36) = 1.23$, $P = 0.32$].

The relationship between the PSE for loudness, the τ_e, and the ϕ_1 is shown in Fig. 5. When the τ_e was between 10 and 100 ms, the perceived loudness increased with increasing τ_e. The ϕ_1 was not correlated with the perceived loudness.

Figure 6 shows the estimated loudness of the IRN stimuli from the models proposed by Moore et al. (1997) and Zwicker and Fastl (1999). When the delay time was 0.5 and 1 ms, the estimated loudness decreased with increasing number of iteration. When the delay time was 2 ms, the estimated loudness decreased slightly with increasing number of iteration. When the delay time was 4, 8, and 16 ms, the estimated loudness did not change as a function of number of iteration. When the delay time was 0.5, 1, 8, and 16 ms, the estimated loudness agreed well with the loudness derived from the present experiment. When the delay time was 2 and 4 ms, the estimated loudness did not correspond to the loudness derived from the present experiment.

4. DISCUSSION AND CONCLUSIONS

When the τ_e was between 10 and 100 ms, the loudness of the IRN increased with increasing τ_e. This is consistent with previous findings using BN (Sato et al. 2002; Soeta et al. 2004) and IRN (Soeta et al. 2007), clearly confirming that loudness is influenced by the repetitive components of sounds in the τ_e range between 10 and 100 ms. Therefore, the τ_e could be useful criteria for measuring loudness. In addition, we found that the increase in loudness for the τ_e values between 10 and 100 ms was approximately 5 dB.

When the delay of the IRN was 0.5 ms, the loudness of the IRN decreased with increasing τ_e. When the delay of the IRN was 1 ms, the loudness of the IRN decreased with increasing τ_e up to approximately 4 ms, and then increased with in-

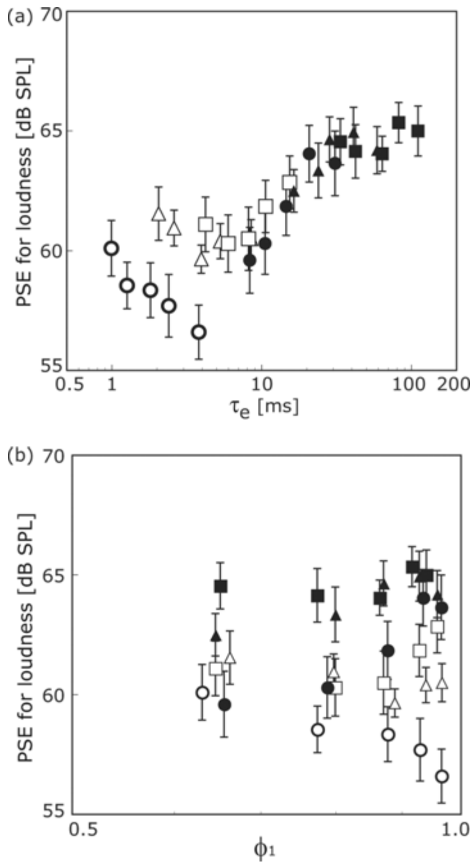

Figure 5. Mean PSE for loudness as a function of the (a) τ_e and (b) ϕ_1 with a delay of (○) 0.5, (△) 1, (□) 2, (●) 4, (▲) 8 and (■) 16 ms. Error bars represent the standard error of the means.

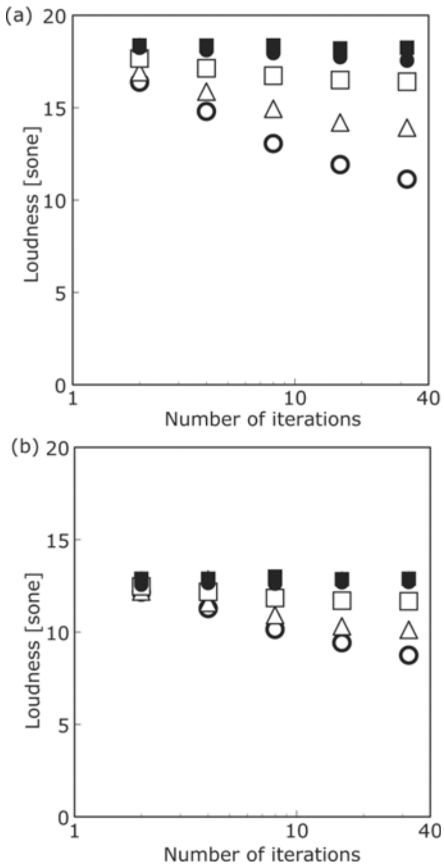

Figure 6. Estimated loudness as a function of the number of iterations with delays of (○) 0.5, (△) 1, (□) 2, (●) 4, (▲) 8 and (■) 16 ms by the models proposed by (a) Moore et al. (1997) and (b) Zwicker and Fastl (1999).

creasing τ_e. These tendencies might be due to the effect of the critical band, that is, loudness remains constant as the bandwidth of the noise increases up to the critical band, and then increases with increasing bandwidth beyond the critical band. This can be predicted by loudness models (Moore et al. 1997; Zwicker and Fastl 1999), as shown in Fig. 6. The power spectrum of the IRN has a ripple with peaks at integer multiples of the reciprocal of the delay, as shown in Fig. 2. Figure 7 shows the bandwidth of the IRN stimuli as a function of the number of iterations. Bandwidth is defined as the bandwidth measured at 3 dB down from the peak. The bandwidth of the IRN stimuli decreases with an increasing number of iterations and the bandwidth of the IRN stimuli at the delay times of 0.5 ms and 1 ms with the number of iterations of 2, 4 and 8 is larger than the critical bandwidth. This suggests that the IRN stimuli at the delay times of 0.5 ms and 1 ms with the number of iterations of 2, 4 and 8 is perceived to be louder than that with the number of iterations of 16 and 32 because of the effect of the critical band. If the bandwidth of the IRN stimuli at the delay times of 0.5 ms and 1 ms with the number of iterations of 2, 4 and 8 was smaller than the critical bandwidth at the reciprocal of the delay, then the loudness increased with an increasing number of iterations, similar to the IRN stimuli at the delay times of 2, 4, and 8 ms. Hence, the effect of the critical band on loudness might impair the effect of the repetitive component on loudness when the IRN stimuli at the delay times of 0.5 ms and 1 ms have 2, 4, and 8 iterations.

When the τ_e was more than 100 ms, the loudness of the IRN remained constant, suggesting the contribution of perceived pitch strength. Previous studies have shown that pitch strength for the IRN is strongest for pitches of approximately 500 Hz, and that stimuli had essentially no pitch strength for pitches below 50 Hz and above 2000 Hz (Yost and Hill 1978). When the τ_e was approximately 100 ms in the present study, the delay time of the IRN was 16 ms and the reciprocal of the delay was 62.5 Hz; this indicates that the perceived pitch was weak. Hence, the loudness might not have increased with increasing τ_e because of the perceived pitch weakness. In addition, the perceived pitch weakness might have also affected the loudness judgment when the delay time of the IRN was 0.5 ms and the reciprocal of the delay was 2000 Hz.

The effect of the τ_e on loudness was not seen in some of the subjects, indicating that there could have been subjects who were hardly affected by the τ_e with respect to loudness judgment. Thus, the effect of the τ_e on loudness judgment seems to partially depend upon the individual. This is consistent

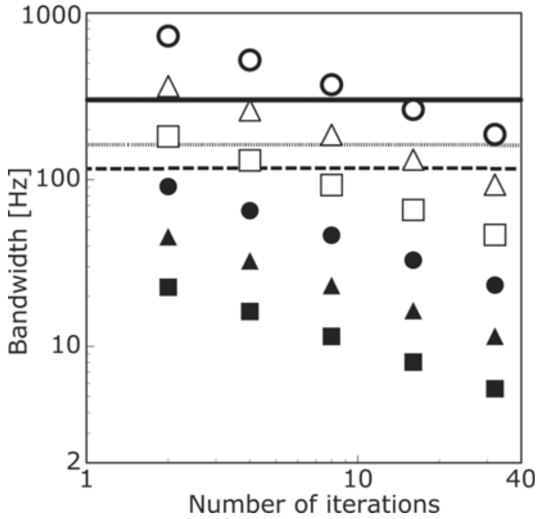

Figure 7. Bandwidths of the IRN stimuli used in the experiment as a function of the number of iterations with delays of (○) 0.5, (△) 1, (□) 2, (●) 4, (▲) 8 and (■) 16 ms. Bandwidth is defined as the bandwidth measured at 3 dB down from the peak at the reciprocal of the delay. Solid, short-dashed and long-dashed lines indicate critical bandwidth at a frequency of 2000, 1000 and 500 Hz, respectively (Zwicker 1961).

with previous findings using BN (Sato et al. 2002; Soeta et al. 2004) and IRN (Soeta et al. 2007).

The previously introduced loudness model (Moore et al. 1997; Zwicker and Fastl 1999) cannot predict loudness when the delay is 2 or 4 ms, which corresponds to a pitch of 500 and 250 Hz, respectively. This suggests the τ_e is useful for supplementing the loudness model. Loudness increases caused by a tonal component might be predictable by the τ_e in a certain range. The τ_e of various noise sources, such as airplanes (Fujii et al. 2001; Sakai et al. 2001), trains (Sakai et al. 2002), motor bikes (Fujii et al. 2002) and flushing toilets (Kitamura et al. 2002) have been analyzed, with results indicating that the τ_e varied within a range of 1-200 ms according to the type of noise source. We found that the increase in loudness for the τ_e values from 10 and 100 ms correspond to an approximately 5 dB increase in SPL. Hence, τ_e values extracted from the ACF should be useful for supplementing the current loudness model when the τ_e value is between 10 and 100 ms.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Young Scientists (A) from the Japan Society for the Promotion of Science (18680025).

REFERENCES

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

Ando, Y., A theory of primary sensations and spatial sensations measuring environmental noise. J. Sound and Vib. 241, 3-18 (2001).

Fraisse, P., Perception and estimation of time. Annu. Rev. Psychol. 35, 1-36 (1984).

Fujii, K., Soeta, Y., and Ando, Y., Acoustical properties of aircraft noise measured by temporal and spatial factors. J. Sound and Vib. 241, 69-78 (2001).

Fujii, K., Atagi, J., and Ando, Y., Temporal and spatial factors of traffic noise and its annoyance. J. Temporal Des. Arch. Environ. 2, 33-41 (2002). http://www.jtdweb.org/

Greenwood, D.D., Auditory masking and critical band. J. Acoust. Soc. Am. 33, 484-502 (1961).

Kitamura, T., Shimokura, R., Sato, S., and Ando, Y., Measurement of temporal and spatial factors of a flushing toilet noise in a downstairs bedroom. J. Temporal Des. Arch. Environ. 2, 13-19 (2002). http://www.jtdweb.org/

Levitt, H., Transformed up-down procedures in psychophysics. J. Acoust. Soc. Am. 49, 467-477 (1971).

Moore, B.C.J., Glasberg, B.R., and Baer, T., A model for the prediction of thresholds, loudness, and partial loudness. J. Audio Eng. Soc. 45, 224-240 (1997).

Moore, B.C.J., Vickers, D., Baer, T., and Launer, S., Factors affecting the loudness of modulated sounds. J. Acoust. Soc. Am. 105, 2757-2772 (1999).

Sakai, H., Sato, S., Prodi, N., Pompoli, R., and Ando, Y., Measurement of regional environmental noise by use of a PC-based system. An application to the noise near the airport "G. Marconi" in Bologna. J. Sound and Vib. 241, 57-68 (2001).

Sakai, H., Hotehama, T. Prodi, N., Pompoli, R., and Ando, Y., Diagnostic system based on the human auditory-brain model for measuring environmental noise -An application to the railway noise-. J. Sound and Vib. 250, 9-21 (2002).

Sato, S., Kitamura, T., and Ando, Y., Loudness of sharply (2068 dB/Octave) filtered noises in relation to the factors extracted from the autocorrelation function. J. Sound and Vib. 250, 47-52 (2002).

Scharf, B., Loudness summation and spectrum shape. J. Acoust. Soc. Am. 34, 228-233 (1962).

Soeta, Y., Maruo, T., and Ando, Y., Annoyance of bandpass filtered noises in relation to the factor extracted from autocorrelation function. J. Acoust. Soc. Am. 116, 3275-3278 (2004).

Soeta, Y., Yanai, K., Nakagawa, S., Kotani, K., and Horii, K., Loudness in relation to iterated rippled noise. J. Sound and Vib. 304, 415-419 (2007).

Yost, W.A., and Hill, R., Strength of the pitches associated with ripple noise. J. Acoust. Soc. Am. 64, 485-492 (1978).

Yost, W.A., Pitch strength of iterated ripple noise. J. Acoust. Soc. Am. 100, 3329-3335 (1996).

Yost, W. A., Fundamentals of hearing; An introduction, Academic press, San Diego, 2000.

Zhang, C., and Zeng, F.G., Loudness of dynamic stimuli in acoustic and electric hearing. J. Acoust. Soc. Am. 102, 2925-2934 (1997).

Zwicker, E. and Fastl, H. Psychoacoustics: Facts and Models, Springer, New York, 1999.

Zwicker, E., Flottorp, G., and Stevens, S.S., Critical bandwidth in loudness summation. J. Acoust. Soc. Am. 29, 548-557 (1957).

Zwicker, E., Subdivision of the audible frequency range into critical bands (Frequenzgruppen). J. Acoust. Soc. Am. 33, 248 (1961).