Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Sound and Vibration 287 (2005) 625–636
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Short Communication

A method for prediction of listening score and psychological impression in an actual noise environment

Takahiro Tamesue[a,*], Shizuma Yamaguchi[a], Tetsuro Saeki[a], Yuichi Kato[b]

[a]Faculty of Engineering, Yamaguchi University, 2-6-11 Tokiwadai, Ube 755-8611, Japan
[b]Interdisciplinary Faculty of Science and Engineering, Shimane University, 1060 Nishikawatsu-cho, Matsue 690-8504, Japan

Received 11 November 2003; received in revised form 6 July 2004; accepted 18 February 2005
Available online 27 April 2005

Abstract

The authors discuss a method for estimating and/or predicting the listening score and two psychological impressions of speech audibility and annoyance, when listening to audio signals, in an actual noise environment. The method introduced the instantaneous spectral distance that reflects the relationship between the spectral level of the speech peaks and that of noise within a short time scale. The validity and the applicability of the proposed method were confirmed experimentally, and reasonable results were obtained. Specifically, it is necessary to predict a psychological impression of the noise as a whole over a long period of time, after a certain amount of exposure to noise, taking the period of higher sound pressure level of the noise into consideration.

1. Introduction

One of the most important problems to secure excellent transmission characteristics of speech information and achieve a comfortable sound environment, in a building used by a great variety of different people in a public city space, to say nothing of spaces used for intellectual or mental work such as schools and offices. A method for commonly evaluating the listening score and psychological impressions of an audio signal and noise, has been discussed in the previously

*Corresponding author. Tel.:/fax: +81 26 295 7090.
E-mail address: tamesue@ei.nagano-nct.ac.jp (T. Tamesue).

doi:10.1016/j.jsv.2005.02.025

published [1]. However, this research limited noise to steady noise with no fluctuation in sound pressure level and frequency components. In a real acoustic environment in which audio signals are transmitted and listened to, both the sound pressure level and the frequency component of the noise show an irregular fluctuation over time. Taking the above into consideration, this paper proposes a prediction method for the listening score and psychological impression of an audio signal and noise in an actual noise environment. More specifically, the instantaneous spectral distance I is first introduced. This reflects the mutual relationships between the spectrum level of the audio signal and that of the noise within a short time scale. Next, by using this index, prediction problems of the listening score and two psychological impressions are discussed with regard to how the listening score/psychological impressions improve when noise insulation is carried out. The predicted results are in good agreement with the observed values.

2. Index for evaluating of listening score and psychological impression

In the previous paper [1], the most useful index for evaluating the listening score and the psychological impression of the audio signal and the noise were discussed under the condition of meaningless steady noise while listening to an audio signal composed of monosyllables. As a result, weighted-mean spectral distance WSPD was selected. $f_t(\mathrm{WSPD})$, $f_A(\mathrm{WSPD})$ and $f_F(\mathrm{WSPD})$ that represent the relationships between the listening score/speech audibility/annoyance and weighted-mean spectral distance have been reported, are respectively shown as follows:

$$f_t(\mathrm{WSPD}) = \frac{98.05}{1 + 0.77\mathrm{e}^{-0.13\mathrm{WSPD}}}, \tag{1}$$

$$f_A(\mathrm{WSPD}) = 0.09\mathrm{WSPD} + 2.88, \tag{2}$$

$$f_F(\mathrm{WSPD}) = -0.09\mathrm{WSPD} + 4.9. \tag{3}$$

In this paper, based on the above index, to evaluate the listening score, speech audibility and annoyance when listening to audio signal in an actual noise environment, the instantaneous spectral distance I is introduced. This reflects the relationships between the spectrum level of the audio signal and that of the noise within a short time scale. Instantaneous spectral distance can be calculated as follows:

$$I = \sum_{i=1}^{8} a_i[L_S(f_i) - L_N(f_i)], \tag{4}$$

where a_i denotes the weights considered percentages that 20 frequency bands [2] which contribute equally to speech intelligibility are included in octave bands with center frequency f_i ($f_1 = 63$, $f_2 = 125, \ldots, f_8 = 8000\,\mathrm{Hz}$), are shown as follows:

$$\begin{aligned} a_1 &= 0.000000, & a_2 &= 0.000000,\\ a_3 &= 0.063794, & a_4 &= 0.140096,\\ a_5 &= 0.226255, & a_6 &= 0.319855,\\ a_7 &= 0.227360, & a_8 &= 0.022640 \end{aligned} \tag{5}$$

and, $L_S(f_i)$ denote band level with center frequency f_i ($f_1 = 63, f_2 = 125, \ldots, f_8 = 8000\,\text{Hz}$) of the speech peaks. In this paper, maximum band levels of audio signal, measured with a real-time octave-band analyzer (RION, type SA-30) with FAST dynamic response for 180 s, are adopted as band levels of speech peaks. $L_N(f_i)$ denote band level with center frequency f_i ($f_1 = 63, f_2 = 125, \ldots, f_8 = 8000\,\text{Hz}$) of the noise with in a short time scale, measured with a real-time octave-band analyzer (RION, type SA-30) with FAST dynamic response, sampling frequency 10 Hz. Instantaneous spectral distance can be only calculated by Eq. (4) when the overall value of band level of speech peaks is between 62 dB and a low value 30 dB from it.

3. Outline of psychological listening experiment

The psychological listening experiment was conducted to compare the observed values of the listening score, and the two psychological impressions of the noise and audio signal with the predicted values. This experiment was as follows.

3.1. Location

The psychological listening experiment was conducted in a simple soundproof room on campus having the following dimensions: length 5.1 m, width 3.3 m, and height 2.2 m. The sound pressure level of the background noise was about 37 dB. A-weighted sound pressure level was about 21 dB (A). The sound pressure level in this paper is the value measured by a sound level meter with FLAT response. The reverberation time was 0.08, 0.07, and 0.06 s for octave band-limited white noise with center frequencies of 500, 1000, and 2000 Hz. It is thought that room reflections have affected neither an audio signal nor noise greatly.

3.2. Subjects

A total of 96 people, 82 male and 14 female students, all with normal hearing participated in the psychological experiment.

3.3. Presented sound

3.3.1. Audio signal

As same as the previous paper [1], A monosyllable list (a list contained 50 monosyllables) from "CD for the evaluation of fitting condition with hearing aids (TY-89)" [3] was used. The overall sound pressure level of the speech peaks was about 62 dB. The band levels of speech peaks are shown in Fig. 1.

3.3.2. External noise

The following noises were adopted as examples of a typical irregular fluctuating noise.

(a) *Non-stationary road traffic noise*: Actual road traffic noise under interrupted traffic flow conditions, recorded in advance over about 2 h at the side of the road. The equivalent continuous sound pressure level was adjusted 62 and 67 dB.

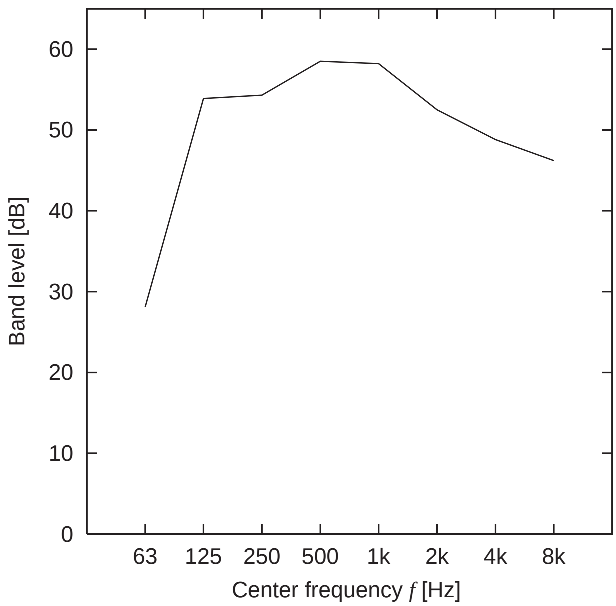

Fig. 1. Band level of speech peaks.

(b) *Stationary road traffic noise*: Actual road traffic noise under uninterrupted traffic flow conditions obtained from "Audio/Acoustics Technical CD for professional use". The equivalent continuous sound pressure level was adjusted 62 dB.

(c) *Aircraft noise*: Actual aircraft noise during take off obtained from "Audio/Acoustics Technical CD for professional use". The equivalent continuous sound pressure level was adjusted 62 and 67 dB.

The equivalent continuous sound pressure level of traffic noise measured at the place contiguous the road is normally higher than 67 or 62 dB, but the lower sound pressure level is also shown due to distance from the road or anything in actual noise environment. Besides, when the equivalent continuous sound pressure level of the noise was set as higher level than presented level in this paper, the psychological listening experiment could not be conducted because an audio signal could not be heard hardly. Therefore, the sound pressure level of noise was set as a value that an audio signal could be heard. Moreover, in order to obtain various listening scores and psychological evaluations, some kinds of noise which have different power spectrum and levels were used. Thus, the equivalent continuous sound pressure level of the noise has been set to disperse the observed values of listening score and psychological evaluation widely. From these reasons, 67 or 62 dB was just used as an example of equivalent continuous sound pressure level of actual noise. The above noise types were assumed to be produced before an insulation wall was installed. Further, as typical examples of noise after installation, these were passed through an RC circuit (see Fig. 2). A transfer function of the RC circuit that represents the insulation wall is shown as follows:

$$G(\mathrm{j}w) = \frac{1}{1 + \mathrm{j}\omega T}. \tag{6}$$

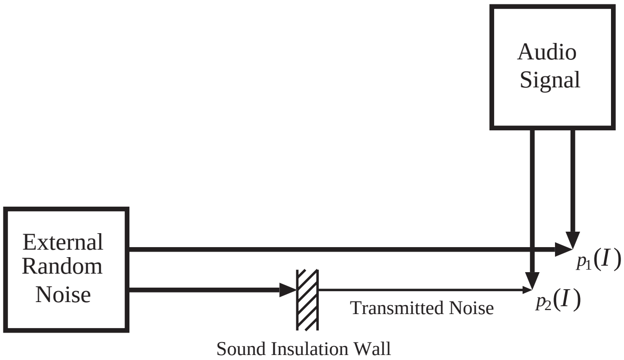

Fig. 2. Relationships between external noise and audio signal before and after installation of sound insulation wall.

Here, a time constant T is 3.8×10^{-3} s. These noises corresponded to transmitted noise normally incident on a single wall with a surface density of about $3.2\,\mathrm{kg/m^2}$.

In addition, the following condition for prediction of the listening score and speech audibility was considered.

(d) No external noise.

3.4. Measurement method of listening score and psychological impressions

The number of subjects for each noise condition was 16, but eight subjects participated in the psychological listening experiment simultaneously. Both the audio signal and the noise were transmitted from the speakers to all of eight subjects seated about 2.2 m from the speakers. It was confirmed ahead of time that there was no difference in the sound pressure levels at the subjects' ears. Subjects noted down the monosyllables they heard. The number of correct monosyllables heard was checked and the listening score was defined as a percentage of correct results after listening to all 50 monosyllables. In addition, the psychological impressions of the noise and the audio signal were investigated. To quantify the psychological evaluation of the noise, various psychological evaluation scales for external noise were possible. Here, as same as previous paper [1], we adopted the seven categorized psychological impressions F_i $(i = 1, 2, \ldots, 7)$ proposed by Furihata [4] (F_1: not at all annoying, F_2: not annoying, F_3: not too annoying, F_4: slightly annoying, F_5: annoying, F_6: very annoying, F_7: extremely annoying). On the other hand, the scale for psychological evaluation of the audio signal adopted the following seven categorized psychological impressions A_i $(i = 1, 2, \ldots, 7)$ of speech audibility [5] (A_1: very inaudible, A_2: quite inaudible, A_3: slightly inaudible, A_4: medium, A_5: slightly audible, A_6: quite audible, A_7: very audible). Concretely, subjects were asked as follows: "After listening to 50 monosyllables, select one from seven categorized psychological impression for annoyance of the noise. Further, select one from seven categorized psychological impression for speech audibility, but judge these two psychological impressions as independently as possible, and judge by your self to the last." The above two psychological evaluations F_i $(i = 1, 2, \ldots, 7)$ of the noise and A_i $(i = 1, 2, \ldots, 7)$ of the audio signal were completed. This operation was carried out with the same subjects for an external noise condition.

4. Prediction of listening score and psychological impression

When the sound pressure level and frequency component of noise show an irregular fluctuation over time, instantaneous spectral distance is a random variable. If a regression models $f_t(I)$, $f_A(I)$, $f_F(I)$ to the listening score, speech audibility and annoyance based on instantaneous spectral distance, and a probability density function $p(I)$ on instantaneous spectral distance are known, employing the regression model and the probability density function, averages of the listening score and psychological impressions of the audio signal and the noise can be calculated as follows:

$$\langle * \rangle = \int_D f_*(I)p(I)\,\mathrm{d}I \quad (D = [-40, 40]\,\mathrm{dB}), \tag{7}$$

where $*$ denotes t, A or F in the case of listening score, speech audibility, or annoyance. In this paper, from a practical point of view, Eqs. (1)–(3) was adopted as $f_t(I)$, $f_A(I)$, and $f_F(I)$.

4.1. Before installation of the sound insulation wall

For the purpose of predicting the listening score and psychological impression for each presented sound by using Eq. (7), it is necessary to provide their probability density function on instantaneous spectral distance previously. The probability distribution was obtained for each presented sound of the psychological listening experiment. The example of the result in the case of noise (a) (67 dB) is shown in Fig. 3. Black marks in Fig. 4 show the comparisons between the predicted values of the listening score obtained from Eq. (7), and the observed values obtained directly from the measured data for each presented sound of the psychological listening experiment. The following findings are revealed by the figure: in spite of predicting the listening score from the psychological listening experiment in which other subjects participated, in the case of irregularly fluctuating noise with various power spectral level forms, a high level of consistency is seen between the predicted and observed values.

Next, the psychological impressions of the audio signal and the noise were predicted. Black marks in Figs. 5 and 6 show, respectively, the comparisons between the predicted values of speech audibility and annoyance, and the observed values. Fig. 5 shows a high level of consistency between the predicted and observed values except the case (c) (62 dB). On the other hand, in Fig. 6, regarding annoyance, the predicted values were lower than the observed values for each presented sound. This tendency appears especially in the case of (a) (67 dB), (c) (62, 67 dB) where the noise had great fluctuations in the sound pressure level and the frequency component. The results indicate that the psychological impression of the noise as a whole, over a long period of time, was larger than the average values for psychological impression over a short time scale, while following to a certain extent the changes in the noise. It was decided that this was because subjects kept not only the psychological impression over a short time scale, but a psychological impression based on memory after a certain amount of exposure to noise, and the psychological impression of the periods with higher noise sound pressure level noise was strongly retained. Though a consideration that does not contradict the above results has already been reported [6], this raises the possibility of bringing about different results depending on the kind of sound source, the number of occurrences, and the presentation order. Taking this point of view, using

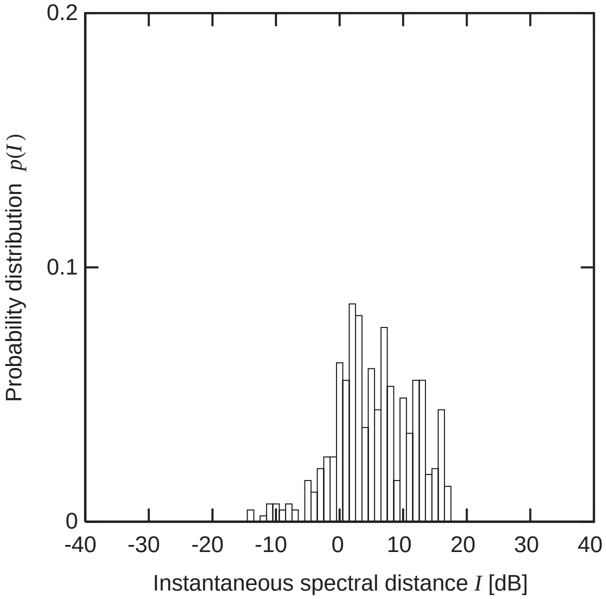

Fig. 3. Probability distribution on instantaneous spectral distance ((a) 67 dB, before installation).

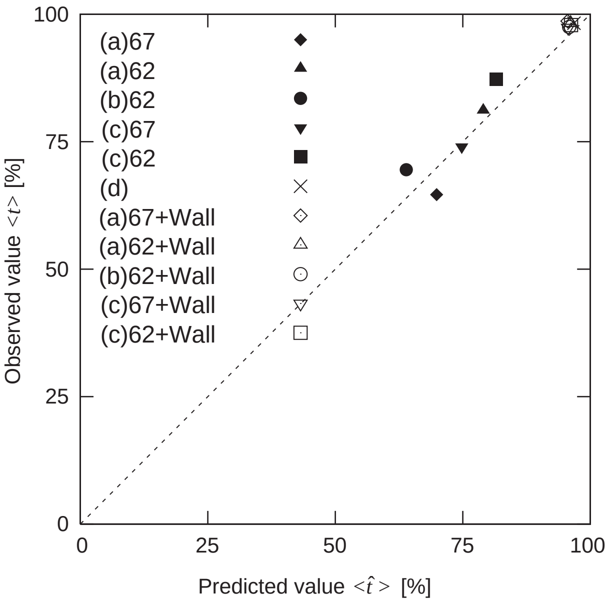

Fig. 4. Comparisons between predicted and observed values for listening score (black marks: before installation, white marks: after installation).

the instantaneous spectral distance in the period of higher noise sound pressure level, the probability distribution was newly set up for each presented sound of the psychological listening experiment. Concretely, the range is assumed to be between the value of the maximum sound

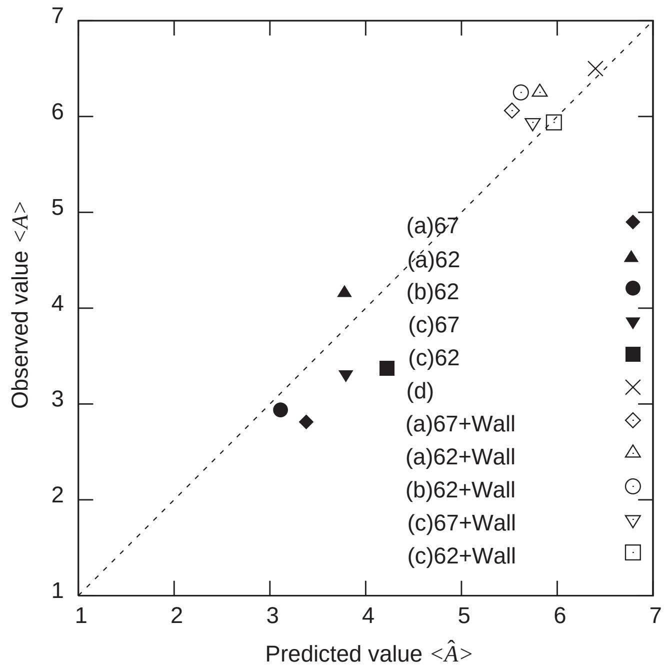

Fig. 5. Comparisons between predicted and observed values for speech audibility (black marks: before installation, white marks: after installation).

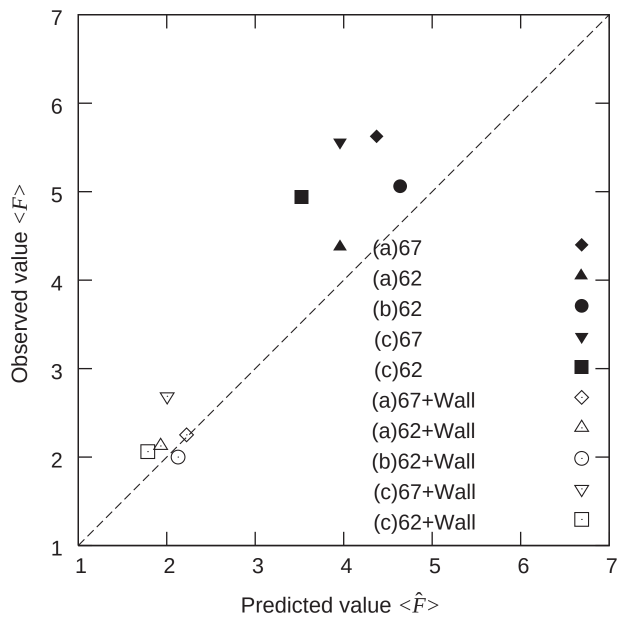

Fig. 6. Comparisons between predicted and observed values for annoyance (black marks: before installation, white marks: after installation).

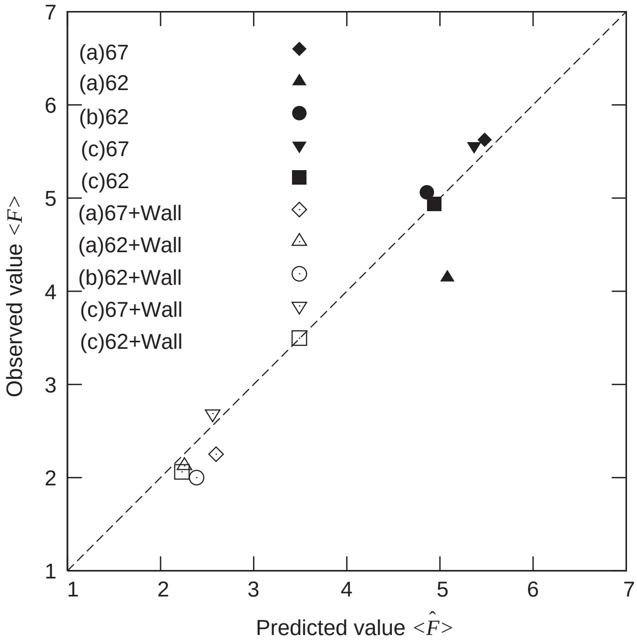

Fig. 7. Comparisons between predicted and observed values for annoyance ($L_N \geqslant L_{\mathrm{Nmax}} - 10$, black marks: before installation, white marks: after installation).

pressure level of the noise and a low value 10, 20 and 30 dB from it. The prediction results in the case that the range is assumed to be between the value of the maximum sound pressure level of the noise and a low value 10, 20 and 30 dB from it, are shown in Figs. 7–9, respectively. Black marks show the comparisons between the predicted values of the annoyance obtained from Eq. (7), and the observed values. The predicted values are in good agreement with the observed results in Fig. 7. In contrast, poor results are obtained in Figs. 8 and 9. Consequently, it turns out that the highest prediction accuracy is acquired when the range isassumed to be between the value of the maximum sound pressure level of the noise and a low value 10 dB from it. It should be consider in more detail how annoyance is determined by the period of the higher sound pressure level of noise furthermore. Based on the above result, it is necessary to predict a psychological impression of the noise as a whole over a long period of time after a certain amount of exposure to noise, taking the period of higher noise sound pressure level into consideration.

4.2. After installation of the sound insulation wall

The acoustical characteristics of noise change when installing things such as sound insulation walls against fluctuating noise. Therefore, probability density function on instantaneous spectral distance after installation of the sound insulation wall is different from that before installation (see Fig. 2). The probability distribution after installation was obtained for the presented sound of the psychological listening experiment. The example of the result in the case of noise (a) (67 dB) is shown in Fig. 10. Compared with Fig. 3 previously shown, probability distribution of the

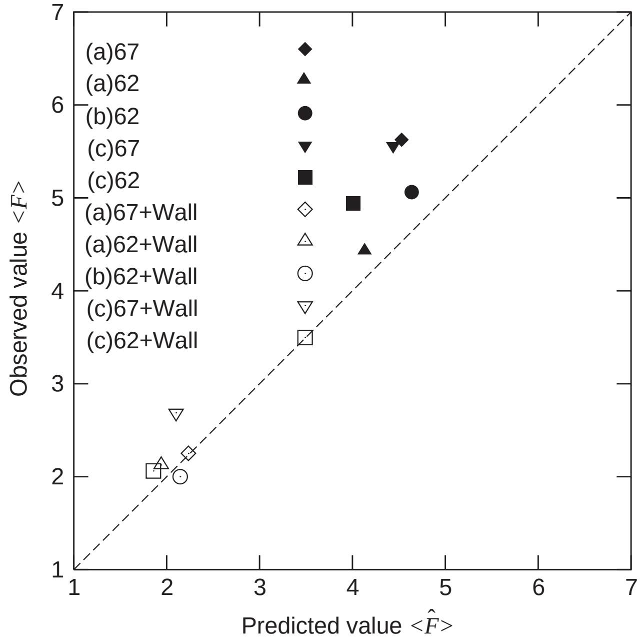

Fig. 8. Comparisons between predicted and observed values for annoyance ($L_N \geqslant L_{\mathrm{Nmax}} - 20$, black marks: before installation, white marks: after installation).

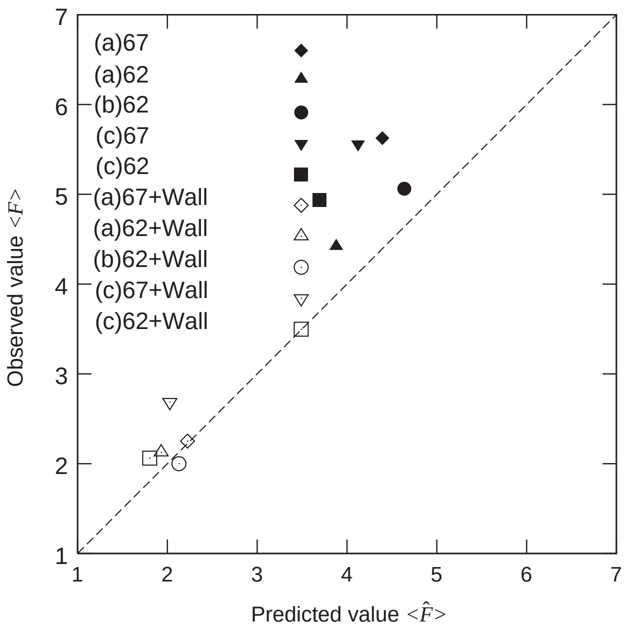

Fig. 9. Comparisons between predicted and observed values for annoyance ($L_N \geqslant L_{\mathrm{Nmax}} - 30$, black marks: before installation, white marks: after installation).

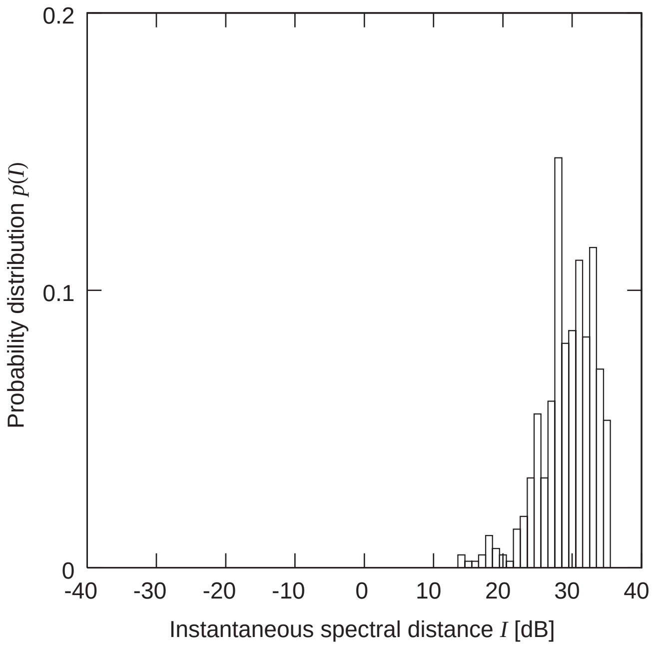

Fig. 10. Probability distribution on instantaneous spectral distance ((a) 67 dB, after installation).

instantaneous spectral distance is translated in the direction of a higher level, and the spread of the values of instantaneous spectral distance is relatively small. White marks in Fig. 4 show the comparisons between the predicted values of the listening score obtained from Eq. (7), and the observed values for each presented sound of the psychological listening experiment after installation. In Fig. 4, regarding the listening score, a higher value than 95% was predicted for each presented sound, in good agreement with the observed values. After the installation, the listening score improved by about 15–40%. White marks in Figs. 5 and 6 show the comparisons between the predicted values of speech audibility and annoyance, and the observed values. With respect to annoyance, in Fig. 6, a good congruence can be seen between the predicted values and the observed values. The reason for the above results is as follows. Even an incident noise on a sound insulation wall has periods of higher sound pressure level and a large range of level fluctuations. In the transmitted noise passed through the sound insulation wall, there is a tendency for the sound pressure level to greatly decrease, and for the range of level fluctuations to become small. Accordingly, it is thought that not appearance of period of the higher sound pressure level which strongly remains in memory as shown in the case of the psychological impression before installation. In the same way, when taking the period of higher noise sound pressure level into account as described in 4.1, annoyance was predicted. White marks in Figs. 7–9 show the result of comparison between the predicted and the observed values. From the Figs. 5 and 7, the prediction results of improvement by 2–3 categories in speech audibility, and decrease by 2–3 categories in annoyance, were in good agreement with observed values.

5. Conclusion

This paper proposed a method for prediction of the listening score and the psychological impressions of an audio signal and a noise, when listening to audio signals in an actual noise environment. The method introduced the instantaneous spectral distance that reflects the relationship between the spectral level of the speech peaks and that of noise within a short time scale. The validity and the applicability of the proposed method were confirmed experimentally, and reasonable results were obtained. Specifically, it is necessary to predict a psychological impression of the noise as a whole over a long period of time, after a certain amount of exposure to noise, taking the period of higher sound pressure level of the noise into consideration.

The main problems to be examined in future are as follows:

(1) Consideration should be given to a situation in which the range of sound pressure level is set, in order to predict the psychological impression of the noise.
(2) The applicability of other prediction methods introducing in theory the probability density function on instantaneous spectral distance should be considered.
(3) The audio signal used in this paper employs monosyllables. However, it is necessary to study cases with more realistic audio signals such as those containing 2 or 3 syllable words, and sentences.

Acknowledgements

The authors would like to express their cordial thanks to Shohei Nishioka and Makoto Kamijo for their helpful assistance. This study was partially supported by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C), No. 11832016, 1999.

References

[1] T. Tamesue, S. Yamaguchi, T. Saeki, Psychological impressions and listening score when listening to audio signals composed of monosyllables and words, while subject to meaningless steady noise—introduction of weighted-mean spectral distance, *Journal of Sound and Vibration* 271 (3–5) (2004) 519–533.
[2] K.D. Kryter, Method for the calculation and use of the articulation index, *The Journal of the Acoustical Society of America* 34 (11) (1962) 1692–1697.
[3] K. Yonemoto, Characteristics of CD for the evaluation of fitting condition with hearing aids (TY-89), *Journal of Otolaryngology Head and Neck Surgery* 11 (9) (1995) 1395–1401 (in Japanese).
[4] K. Furihata, T. Yanagisawa, Investigation on composition of a rating scale possible common to evaluate psychological effects on various kinds of noise sources, *The Journal of the Acoustical Society of Japan* 45 (8) (1989) 577–582 (in Japanese).
[5] T. Nakajima, S. Maeda, The application of speech transmission index (STI) as a measure of Japanese speech audibility, *Proceedings of the Research Committee Meeting on Architectural Acoustics of the Acoustical Society of Japan*, AA-84-30, 1984, pp. 1–8 (in Japanese).
[6] S. Namba, S. Kuwano, T. Kato, J. Hellbruck, Memory of sound level pattern of various noises, *Proceeding of the Fall Meeting of the Acoustical Society of Japan*, 2001, pp. 769–770 (in Japanese).