Journal of Sound and Vibration (2001) **245**(2), 205–215
doi:10.1006/jsvi.2001.3574, available online at http://www.idealibrary.com on IDEAL®

PSYCHOLOGICAL EVALUATION OF EXTERNAL NOISE IN THE CASE OF LISTENING TO AN AUDIO SIGNAL, TAKING ACCOUNT OF THE DIFFERENCE BETWEEN THE POWER SPECTRAL CHARACTERISTIC OF THE AUDIO SIGNAL AND THAT OF NOISE

S. YAMAGUCHI, T. SAEKI, T. TAMESUE AND Y. KATO

Faculty of Engineering, Yamaguchi University, 2-16-1 Tokiwadai, Ube 755–8611, Japan and Interdisciplinary Faculty of Science and Engineering, Shimane University, 1060 Nishikawatsu-Cho, Matsue 690–8504, Japan

(*Received* 10 *August* 2000, *and in final form* 10 *August* 2000)

It is very important to understand quantitatively the relationship between the difference of an audio signal from the external noise and the psychological impression of the noise, to design a comfortable sound environment. In this paper, a method for predicting and/or estimating the human psychological response to the external noise in the case of listening to the audio signal, is proposed by using the fuzzy set theory. Concretely, when the persons listening to the audio signal are exposed to meaningless random noise, bi-variate membership functions, the supports of which are the spectral distance measure and signal-to-noise ratio, are first established. Next, a method for evaluating the psychological response is proposed by introducing the concept of the fuzzy probability. Finally, the validity and the applicability of the proposed method are confirmed experimentally by applying to the actually observed data. The experimental results are in good agreement with the theory.

1. INTRODUCTION

The direct transmission of a speech is one of the most fundamental methods of communication. Realization of a comfortable sound environment in which the listener can concentrate on the speech without being distracted by external noise, is important. It is very important to understand quantitatively the relationship between the difference of the audio signal from the external noise and the psychological impression of the noise, to design a comfortable sound environment. However, the psychological response of listeners to external noise is strictly based on a subjective judgment as a human being. The ambiguity of language expression should be considered fundamental to sensory response, and the ambiguity of the subjective judgment of the human listener is unavoidable.

From these viewpoints, this paper is a further evolution of the discussions already published [1, 2]. The psychological influence of non-white meaningless random noise on humans engaged in psychological activities, such as listening to a lecture, is examined using fuzzy set theory. That is, the difference between power spectral level forms of the speech and meaningless random noise is reflected to membership functions of the psychological impression of the noise. A systematic method for predicting and/or estimating the human psychological response to the meaningless random noise, in the case of the speech and/or noise with arbitrary power spectral level forms, is proposed. Specifically, bi-variate membership functions of the psychological impression ‘annoyance’ [3], the supports of which are the spectral distance measure (the difference between power spectral characteristic of the audio signal and that of the noise) and signal-to-noise ratio, are first established. Next, the psychological response of the listeners to the meaningless random noise with an arbitrary power spectral level form, in the case of listening to a speech by a male or a female voice, is predicted by introducing the concept of the fuzzy probability using the above bi-variate membership functions, which are set up beforehand. Finally, the predicted value and the measured data obtained in the psychological experiment are compared, and fairly good agreement was observed between the predicted value and the measured data.

2. OUTLINE OF THE PSYCHOLOGICAL EXPERIMENT

An outline of the indoor psychological experiment that was conducted is shown below.

2.1. EXPERIMENT I

In order to establish the bi-variate membership functions of the psychological impression, Experiment I was conducted.

[I-A] Location of experiment. A laboratory on campus having the following dimensions: length 7·3 m, width 12·1 m, and height 2·6 m. The sound pressure level of the background noise was in the range of approximately 44–45 dB.

[I-B] Time and date of experiment. 10:00 a.m.–11:00 a.m. and 5:00 p.m.–9:00 p.m. in the last part of September.

[I-C] Subjects. A total of five people, three male students and two female students, with normal hearing.

[I-D] Presented sound

(I-D-1) Speech. A commercially available tape of lecture, by a male voice (lecture time: 80 min). The energy-mean value of sound pressure level is about 63 dB. Sounds on the tapes other than voices, such as applause/sound effects/music, were removed. One sample path of speech is shown in Figure 1(a).

(I-D-2) External noise. The white noise is passed through an octave band-path filter, and then is amplified/attenuated by statistically independent random variables according to a uniform distribution (duration: 10 s). Eight octave filters with center frequencies 63, $125, \ldots, 8000$ Hz are used. The gain is adjusted such that the level variation range of the noises is within 43–90 dB. One sample path of external noise is shown in Figure 1(b).

[I-E] Measurement method of psychological response. Figure 2 shows the setup of the psychological experiment. Five subjects participated in the psychological experiment simultaneously. The speech and the external noise were radiated from speakers to the subjects simultaneously, and the psychological response of the subjects to the noise was investigated. The band-pass filters shown in the figure were used to remove components other than the frequency band-width [45, 11200] Hz for both the speech and the external noise. It was confirmed in advance that there was no difference in the radiation sound

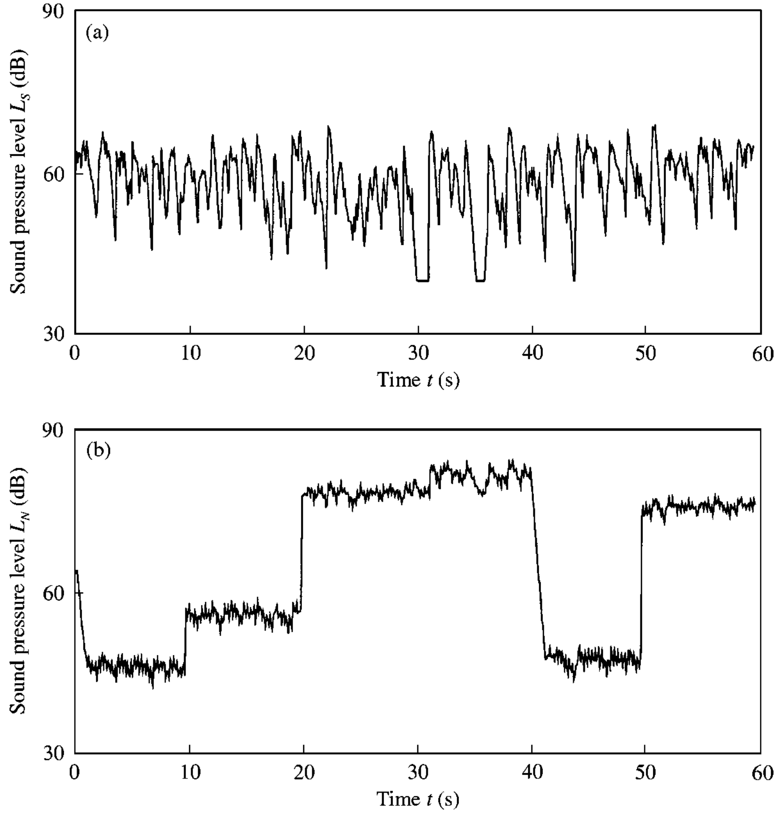

Figure 1. Sample paths of speech and external noise. (a) speech, (b) external noise.

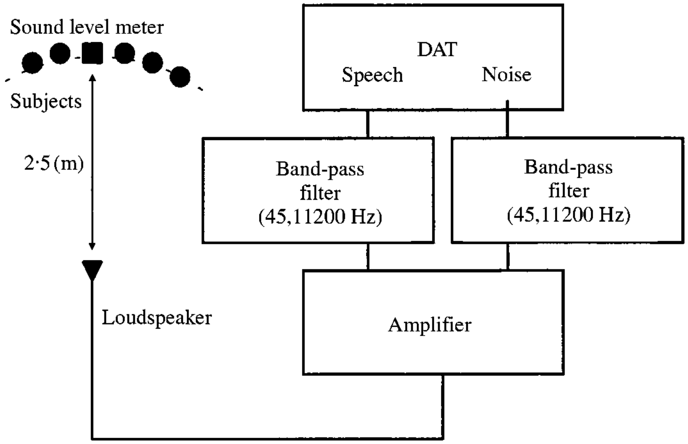

Figure 2. Setup of the psychological experiment.

pressure level from the speakers with respect to the position of the subjects. We adopted the following seven-category evaluation scale Fi ($i = 1, 2, \ldots, 7$) of annoyance [3]:

$F1$: Not at all annoying.

$F2$: Not annoying.

*F*3: Not too annoying.
*F*4: Slightly annoying.
*F*5: Annoying.
*F*6: Very annoying.
*F*7: Extremely annoying.

Subjects listened to the speech and the noise simultaneously, judged the above psychological evaluation Fi ($i = 1, 2, \ldots, 7$) of the noise every 10 s and input the result into a computer by pushing buttons. A large clock with a second hand is placed where all subjects can see it easily. The clock does not affect the acoustic propagation characteristic between the speaker and the subjects so that the time passage of 10 s can be seen prior to the decision of the psychological impression. In addition to this clock, a lamp placed near each subject is lit after 10 s so that the subjects can make the psychological judgment for the prior 10 s immediately after the lighting of the lamp. The sound pressure levels of the speech and the noise at that time were input into a computer simultaneously (sampling frequency: 10 Hz). Subjects were instructed to listen to the speech carefully and to try to understand the content as much as possible. The degree to which each subject was able to understand the content of the speech was not investigated. Before carrying out this experiment, subjects practised beforehand in order to avoid confusion with respect to the judgment and the push button operation. Judgment was continuously recorded every 10 s for 5 min, and these operations were repeated 8 times for a combination of presented sounds while subjects were given sufficient rest in order to avoid fatigue.

2.2. EXPERIMENT II

In order to compare the observed values of psychological response to external noise with the predicted values, Experiment II was conducted.

[II-A] Location of experiment. A simple anechoic room on campus having the following dimensions: length 4·6 m, width 4·5 m, and height 2·6 m. The sound pressure level of the background noise was in the range of approximately 32–37 dB.

[II-B] Time and date of experiment. 9:00 a.m.–6:00 p.m. in the last part of December.

[II-C] Subjects. A total of eight people, six male students and two female students, with normal hearing.

[II-D] Presented sound

(II-D-1) Speech. Four commercially available tapes of lecture, by male voices (A, B, C, D) and three commercially available tapes of lecture, by female voices (E, F, G).

(II-D-2) External noise. Eight octave band-limited white noises with center frequencies 63, 125, ..., 8000 Hz, similar to I-D-2, are first made. Next, two octave band-limited white noises with different center frequencies (f_{c1}, f_{c2}) are synthesized with power ratio 1 : 1 and then they are amplified/attenuated by statistically independent random variables according to a uniform distribution (duration: 10 s). The gain is adjusted such that the level variation range of the noises is within 43–90 dB. Seven combinations of center frequencies of the octave band-limited white noises are listed below.

$$(f_{c1}, f_{c2}) = (63, 125), (125, 250), (250, 500), (4000, 8000) \text{ (Hz)}. \tag{1}$$

Specifically, we adopted the following seven combinations of the above speeches and external noises at random.

A and (63, 125), B and (500, 1000), C and (1000, 2000), D and (4000, 8000),
E and (125, 250), F and (250, 500), G and (2000, 4000) (Hz).

[II-E] Measurement method of psychological response. Eight subjects were divided into two groups. Four subjects participated in the psychological experiment simultaneously. The speech and the external noise were radiated from speakers to the subjects simultaneously, and the psychological response of the subjects to the noise was investigated. The specific method of the psychological experiment is the same as that used in Experiment I. The operations were repeated two times for a combination of presented sounds.

3. DETERMINATION OF BI-VARIATE MEMBERSHIP FUNCTIONS

Using the measured data of Experiment I, bi-variate membership functions:

$$\mu_{Fi}(L, S) \quad (L \in D_1 = [-40, 40],\ S \in D_2 = [-10, 10];\ i = 1, 2, \ldots, 7),$$

$$L = 10\log_{10}\left(\frac{1}{M}\sum_{K=1}^{M} 10^{L_S(K)/10}\right) - 10\log_{10}\left(\frac{1}{M}\sum_{K=1}^{M} 10^{L_N(K)/10}\right),$$

$$S = \mathrm{sgn}\left[\log_2\frac{f_S}{f_N}\right]\int\left(\log\frac{\Phi_S(\omega)}{\Phi_N(\omega)}\right)\Phi_S(\omega)\,d\omega \tag{2}$$

of the psychological impression Fi ($i = 1, 2, \ldots, 7$) to the external noise were set up, where $L_S(K)$ and $L_N(K)$ are sound pressure level values (sampling frequency: 10 Hz) for each speech and external noise respectively. In the specific calculation of signal-to-noise ratio L, average operation was carried out for the remaining 6 s after removing 1 s of the leading times and 3 s of the trailing times from the 10 s of duration time presented for one psychological judgment, in order to avoid the influence of errors in the transient part, in which the sound pressure level value of the randomly amplified/attenuated external noise changes sharply ($M = 60$ in equation (2)). $\Phi_S(\omega)$ and $\Phi_N(\omega)$ in equation (2) are relative power spectral level forms of speech and external noise, respectively ($\log(\int\Phi_*(\omega)\,d\omega) = 0$; $* = S, N$), and f_S and f_N are dominant frequencies of speech and external noise ($\mathrm{sgn}[\log_2(f_S/f_N)] > 0$ (<0) means that dominant frequency of the speech is higher (lower) than that of the noise). In the specific calculation of spectral distance S, power spectral level forms of the speech and noise obtained for the above 6 s are used. The bi-variate membership functions are set up by the following procedure: The measured data for psychological impression Fi is totaled for each value of signal-to-noise ratio L and spectral distance S. Next, these data are displayed graphically on a computer, and a smoothing operation is performed by watching.

The results of the seven bi-variate membership functions established for each psychological impression Fi ($i = 1, 2, \ldots, 7$) are shown in Figure 3. This figure reveals the following: When the value of signal-to-noise ratio is increased (decreased), needless to say, the psychological response to the noise becomes $F1$: Not at all annoying ($F7$: Extremely annoying). When the value of spectral distance S is decreased (increased) so that the frequency of the noise is higher (lower) than that of the speech, the signal-to-noise ratio of the same psychological impression (for example, $F4$: Slightly annoying) becomes higher (lower), the tendency is that the external noise becomes more (less) annoying. In order to

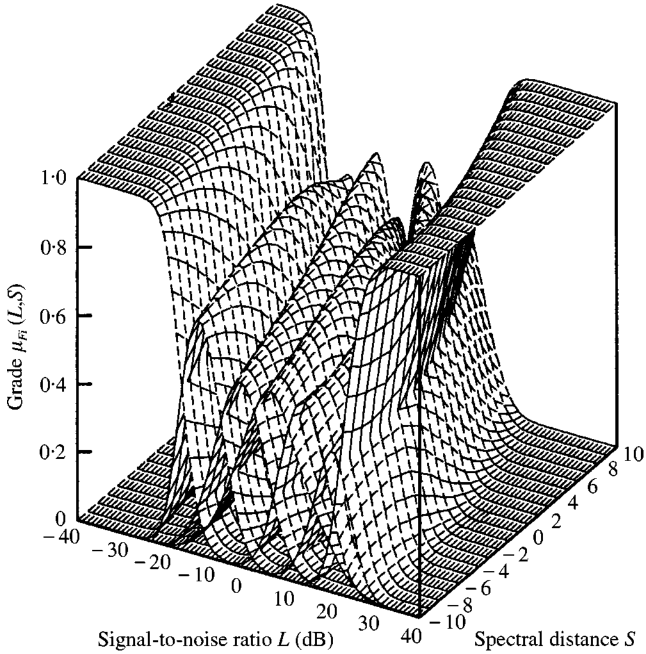

Figure 3. Bi-variate membership functions for psychological impression Fi ($i = 1, 2, \ldots, 7$).

show this fact clearly, membership functions in the case of spectral distance $S = -10, 0, 10$ are shown in Figure 4.

4. PREDICTION OF PSYCHOLOGICAL RESPONSE TO NOISE

As an evaluation method of the psychological response to external noise, we adopted the set of fuzzy probability:

$$P(Fi) = \int_{D_1}\int_{D_2} \mu_{Fi}(L, S)p(L, S)\,\mathrm{d}L\,\mathrm{d}S \quad (i = 1, 2, \ldots, 7), \tag{3}$$

for each psychological impression Fi ($i = 1, 2, \ldots, 7$) as a fuzzy event [4] (time ratio of the occurrence of Fi) and following average psychological impression:

$$\langle F \rangle = \sum_{i=1}^{7} iP(i)\ (P(i) = P(Fi)). \tag{4}$$

Bi-variate probability distributions $p(L, S)$ on signal-to-noise ratio and spectral distance were obtained for each presented sound of Experiment II. As for the example of the result in the case of speech C and noise (1000, 2000) ($[L_1, L_2]$ (the level variation range of L) $= [-40, 40]$) is shown in Figure 5.

The set of predicted time ratio $P(Fi)$ ($i = 1, 2, \ldots, 7$) of the occurrence of each psychological impression Fi obtained from equation (3) using the bi-variate membership functions set up in 3 and Figure 5 ($[L_1, L_2] = [-40{:}40]$) is shown as predicted value in Figure 6. The value of $P(Fi)$ ($i = 1, 2, \ldots, 7$) obtained directly from the measured data of Experiment II is shown as the observed value in the same figure. As for the example of the

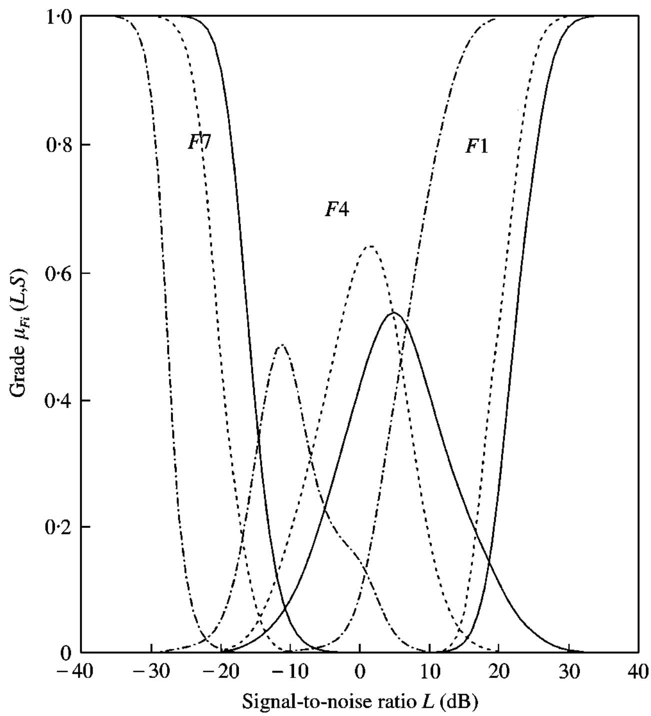

Figure 4. Membership functions $\mu_{Fi}(L, S)$ $(i = 1, 4, 7; S = -10, 0, 10)$. $S = -10$, ——; $S = 0$, ...; $S = 10$, –·–·–.

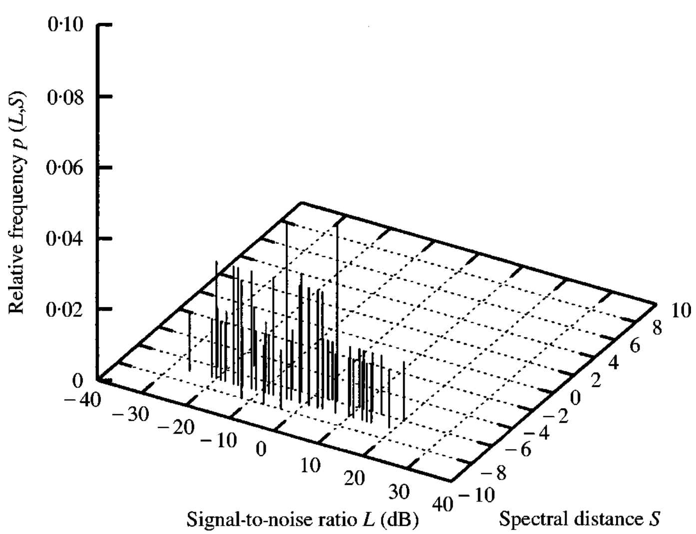

Figure 5. Bi-variate probability distribution $p(L, S)$ (speech C and external noise (1000, 2000); $[L_1, L_2] = [-40, 40]$).

result in the case of speech G and noise (2000, 4000) ($[L_1, L_2] = [-40{:}40]$) is shown in Figure 7 (results for the other cases are omitted). Good agreement is seen between the predicted and observed values. On the other hand, the comparison between the predicted and observed values for the average psychological impression for each of the seven

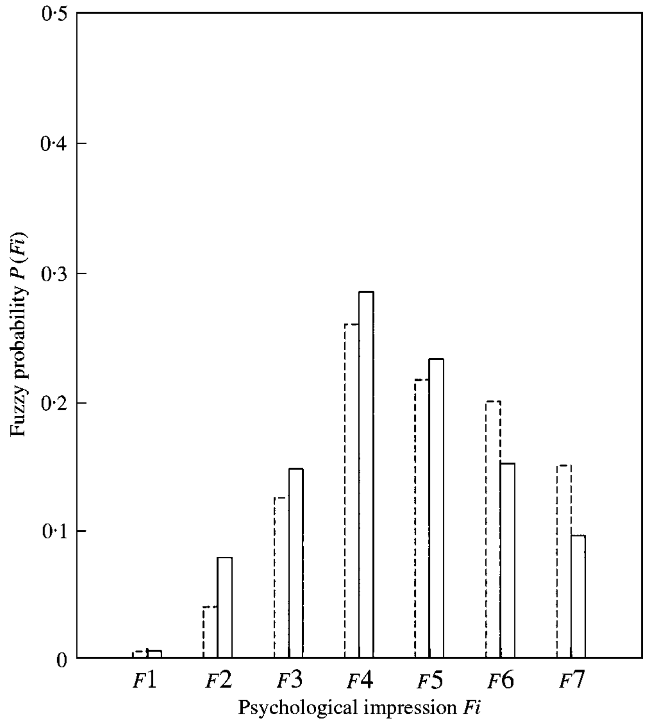

Figure 6. Comparison between predicted and observed values for $P(Fi)$ (speech C and external noise (1000, 2000); $[L_1, L_2] = [-40, 40]$). Observed value, ——; Predicted value, - - - - -.

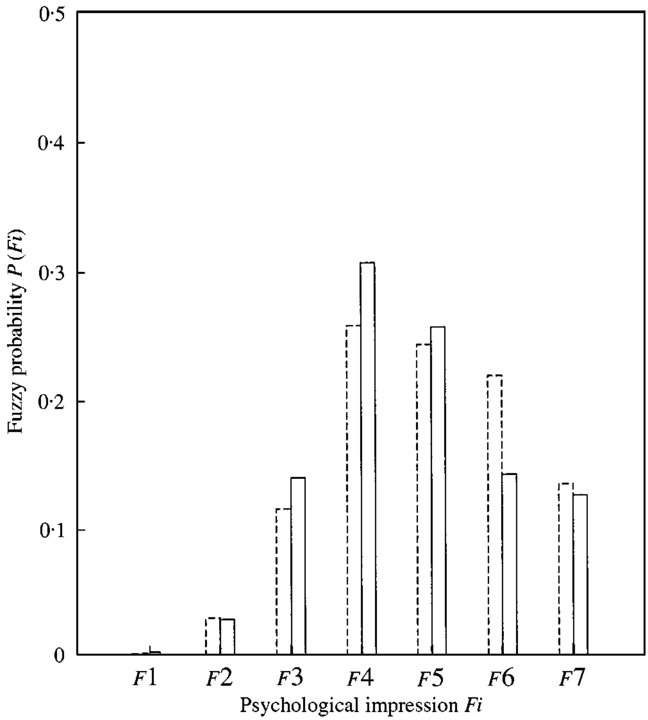

Figure 7. Comparison between predicted and observed values for $P(Fi)$ (speech G and external noise (2000, 4000); $[L_1, L_2] = [-40, 40]$). Observed value, ——; Predicted value, - - - - -.

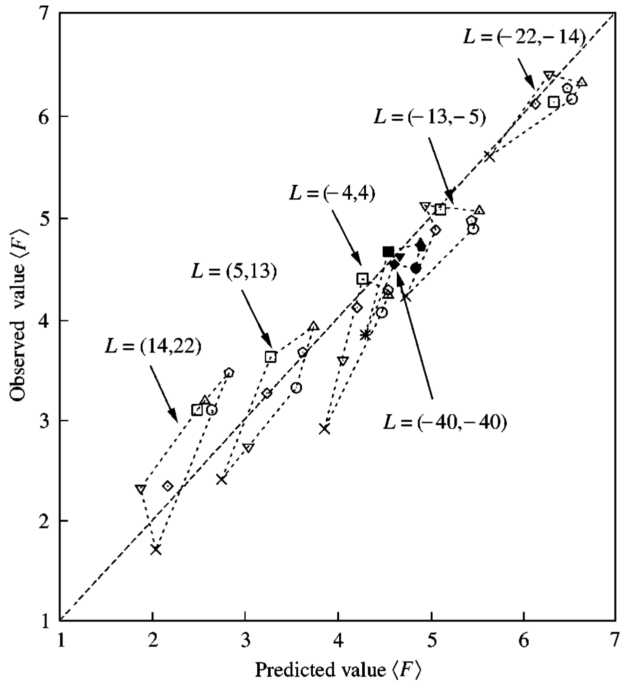

Figure 8. Comparison between predicted and observed values for $\langle F \rangle$. A-(63, 125), ×; B-(500, 1000), ⊡; C-(1000, 2000), ⊙; D-(4000, 8000), △; E-(125, 250), ▽; F-(250, 500), ◇; G-(2000, 4000), ⬠.

combinations of presented sounds ($[L_1, L_2] = [-40:40]$) is shown in Figure 8. The results of various level variation ranges of $L([L_1, L_2] = [-22, 14], [-13, 5], [-4, 4], [5, 13], [14, 22])$ are also shown in the same figure (correlation coefficient $r = 0{\cdot}966$). There is a natural tendency for the external noise to become less annoying as the signal-to-noise ratio value is increased. The following findings are revealed by Figures 6–8: The psychological response of the other subjects can be predicted systematically, by using the bi-variate membership functions proposed in the present paper, in the case of speech and/or meaningless random noise with various power spectral level forms. Therefore, the bi-variate membership functions in Figure 3 are almost stable. From this viewpoint, the psychological impression for various differences of the audio signal from the external noise (signal-to-noise ratio and spectral distance measure) is calculated. The relationship between the average psychological impression $\langle F \rangle$ and signal-to-noise ratio L for each spectral distance $S = -10, -5, 0, 5, 10$, obtained from the bi-variate membership functions in Figure 3, is shown in Figure 9.

5. CONCLUSION

We considered the prediction method of the psychological response to the meaningless random noise in the case of listening to the speech. The psychological response in the case of the speech and/or external noise with arbitrary power spectral level forms can be predicted systematically using bi-variate membership functions on spectral distance (mutual relationship between frequency characteristics of speech and noise) and signal-to-noise ratio (mutual relationship between amplitude characteristics of speech and noise). The validity and the applicability are confirmed experimentally, reasonable results are obtained.

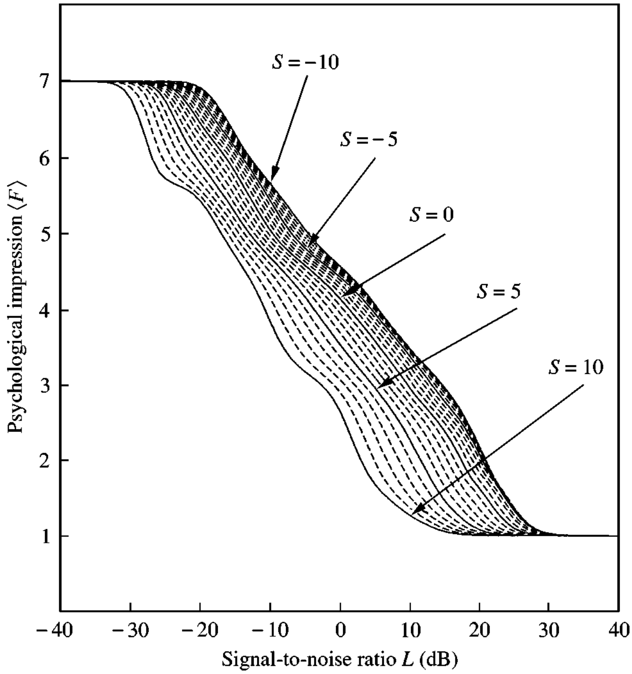

Figure 9. Relationship between psychological impression $\langle F \rangle$ and signal-to-noise ratio L ($S = -10, -5, 0, 5, 10$).

In the present paper, the proof of validity of the evaluation method is emphasized, since the proposed method is in an early stage of study. The primary subjects that should be examined in future studies are listed below:

(1) It is necessary to accumulate the psychological experiment data by many subjects. Especially, it is important to make the bi-variate membership functions which are the basis of the prediction method stable. In order to establish stable membership functions in many situations, it is reasonable to divide the situations of subjects into some categories, such as the temporal factor (daytime, nighttime), the spatial circumstances factor (sickroom, factory, house), and the activity factor (sleep, study, labor).
(2) The prediction method in many cases using actual noises and the case where a pure tone is included in the external noise so that a line spectrum appears in the power spectral level form should be considered experimentally.
(3) The applicability of the same method in the case where the external noise is meaningful noise such as music and conversation should be confirmed.

ACKNOWLEDGMENTS

The authors would like to express their cordial thanks to Atsuko Ohmae and Ryuichi Hirata for their helpful assistance, and to their colleagues who provided valuable comments at The Technical Committee of Sound and Vibration of the Acoustical Society of Japan [5], The 44th Annual Conference of the Institute of Systems Control and Information Engineers [6] and The 41st Annual Conference of the Japan Ergonomics Society [7]. This study was partially supported by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C), No. 11832016, 1999.

REFERENCES

1. S. Yamaguchi, Y. Kato, K. Oimatsu and T. Saeki 1995 *Applied Acoustics* **45**, 139–154. A psychological evaluation method for fluctuating random noise based on fuzzy set theory.
2. S. Yamaguchi, T. Saeki and Y. Kato 1996 *Applied Acoustics* **48**, 155–174. A practical prediction method of psychological response to arbitrary non-white random noise based on simplified patterns of membership functions.
3. K. Furihata and T. Yanagisawa 1989 *The Journal of the Acoustical Society of Japan* **45**, 577–582. Investigation on composition of a rating scale possible common to evaluate psychological effects on various kinds of noise sources (in Japanese).
4. L. A. Zadeh 1968 *Journal of Mathematical Analysis and Application* **23**, 421–427. Probability measures of fuzzy events.
5. S. Yamaguchi, T. Saeki, T. Tamesue and Y. Kato 1999 *Proceedings of the Technical Committee of Sound and Vibration of the Acoustical Society of Japan* **N-99-69**, 1–6. A method for evaluating the psychological response to random noise in the case of listening to audio signal taking account of the difference between power spectral characteristic of audio signal and that of noise (in Japanese).
6. S. Yamaguchi, T. Saeki, T. Tamesue and Y. Kato 2000 *Proceedings of the 44th Annual Conference of the Institute of Systems, Control and Information Engineers* 591–592. A method for evaluating the psychological response to random noise in the case of listening to audio signal taking account of the difference between power spectral characteristic of audio signal and that of noise (in Japanese).
7. S. Yamaguchi, T. Saeki, T. Tamesue and M. Nakamura 2000 *The Japanese Journal of Ergonomics* **36** (Supplement), 370–371. A method for evaluating the psychological response to random noise in the case of listening to audio signal taking account of the difference between power spectral characteristic of audio signal and that of noise (in Japanese).