High directivity masking sound system for achieving speech privacy

Takahiro TAMESUE[1]; Testuro SAEKI[2]

[1] Organization for Academic Information, Yamaguchi University, 2-16-1, Tokiwadai, Ube, Japan

[2] Faculty of Engineering, Yamaguchi University, 2-16-1, Tokiwadai, Ube, Japan

ABSTRACT

In recent years, importance has been attached to achieving speech privacy in open spaces. Generally, although measures such as the use of sound partition are instituted in many cases, measures that use other sounds to mask speech by emitting sound other than speech have also been considered. The masking noise emitted to the area where high level of speech privacy is not required, may cause an increased psychological impression of annoyance, leading to a decline in performance. In this study, we constructed a masking sound system with highly directional sound from modulated ultrasound as a masking noise for achieving speech privacy in the narrow area. Psychological experiments were conducted in which the masking sound was transmitted to participants from frontal or above directions with a high directivity masking sound system. Using the experimental data, the relationships between the degree of speech privacy and frequency characteristics and directivity of a high directivity sound through the masking sound system were investigated.

Keywords: Speech privacy, Masking noise, Parametric speaker array, Psychological evaluaton I-INCE Classification of Subjects Number(s): 63.3
(See `http://www.inceusa.org/links/Subj%20Class%20-%20Formatted.pdf`.)

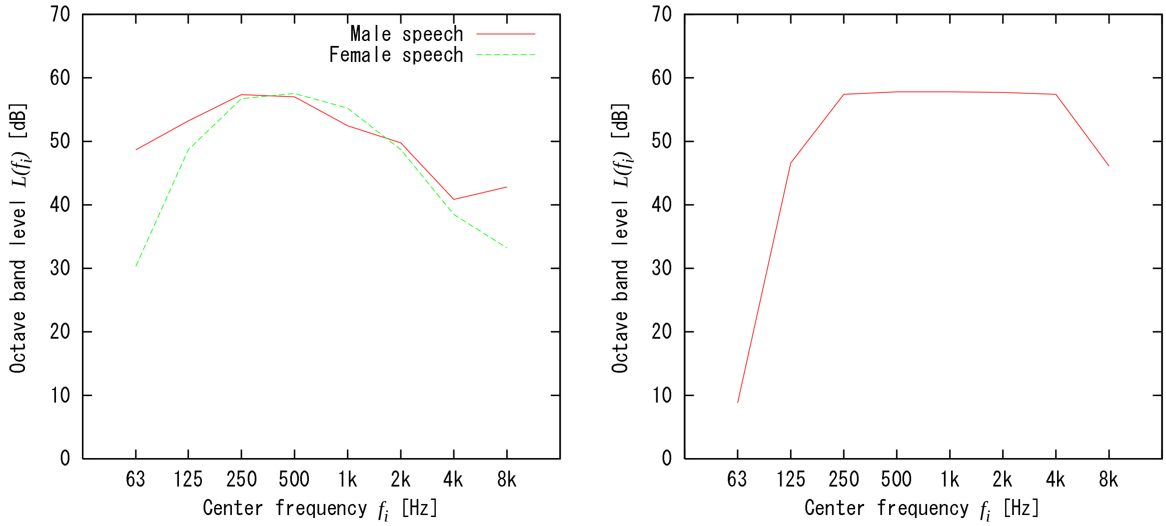
1. INTRODUCTION

In recent years, importance has been attached to achieving speech privacy [1] in open spaces such as for oral consultations near waiting rooms of small-scale clinics, tax-payment consultations at a taxation office window, course consultations in school classrooms, and legal aid services in temporary booths. Generally, although measures such as the use of sound partition are instituted in many cases [2-3], measures that use other sounds to mask speech by emitting sound other than speech have also been considered [4]. A way of masking meaningful speech with meaningless noise would have great benefits. From the above view point, previous studies focused on masking speech with meaningless steady noise for achieving speech privacy achieving [5]. However, the research to date has focused on evaluating speech privacy when the masking noise is emitted from the normal loud speaker system all over the room. The masking noise emitted to the area where high level of speech privacy is not required, may cause an increased psychological impression of annoyance, leading to a decline in performance. In this study, we used a highly directional sound from modulated ultrasound as a masking noise for achieving speech privacy in the narrow area. Psychological experiments were conducted in which the masking sound was transmitted to participants from frontal or above directions with a parametric acoustic array speaker [6]. Using the experimental data, the relationships between the degree of speech privacy and frequency characteristics that directivity of parametric acoustic array speaker were investigated. The results suggested that it is possible to maintain speech privacy in the narrow area by presenting highly directional masking sound.

2. OUTLINE OF PSYCHOPHYSIOLOGICAL EXPERIMENT I

Psychological experiment I was conducted to examine how the psychological impression of speech privacy under the influence of masking noise with a parametric acoustic array speaker. The outline of the indoor psychological experiment was as follows.

[1] tamesue@yamaguchi-u.ac.jp



(A) Speech peaks (B) Masking noise

Figure 1 – Power spectra of speech peaks and masking noise

2.1 Subjects

A total of 16 people, 15 male and 1 female students, all with normal hearing, participated in the psychological experiment I.

2.2 Audio signal

Male speech and female speech, made by deleting handclaps, sound effects, and music, etc.fromcommercially available speech tapes, were used in the experiment. Maximum band levels of speech measured with a sound level meter were adopted as the band levels of speech peaks. The overall sound pressure level ofthe speech peaks was about 62 dB. The power spectra of the speech peaks are shown in Figure 1 (A).

2.3 Masking noise

Band-limited pink noise with [180, 5630] Hz frequency band-width was selected as the best and most-practical meaningless steady noise for robustly masking speech, which has various forms of power spectral levels, at a minimum sound pressure level [7], was used. The sound pressure levels of the masking noise were set at 53, 56, 59, 62, 65, 68, 71 dB. The power spectrum of the masking noise is shown in Figure 1 (B).

2.4 Acoustic loudspeakers

(a) Parametric speaker array

HSS Japan H450 and Tristate K-2617 were used as parametric acoustic speaker array. Figure 2 shows the magnitude frequency response of speakers. These speakers have an output frequency range with a lower limit of approximately 500 Hz. Figure 3 also shows the directivity patterns for parametric speaker array.

(b) Cone loudspeaker

EDIROL MA-10 was employed as a cone loudspeaker for transmitting the audio signal. In addition, it was considered that audio signal was masked with meaningless noise which was presented from cone loudspeaker for comparison of the speech privacy evaluation.

2.5 Measurement

The audio signal was transmitted from frontal direction by a cone loudspeaker and the masking noise was transmitted from frontal or above direction by parametric speaker array to subject. After listening to the speech for 30 s, an evaluation of the speech privacy was made using the following evaluation scales [8]: F_1: Did not notice talking. F_2: Possible to recognize the sound as a voice, though the contents of speech were not understood. F_3: When listened to carefully, the contents of the speech were understood to some extent. F_4: Even if not intensively listened to, all the contents of the speech were understood. The above operations were executed twice in each of the cases.

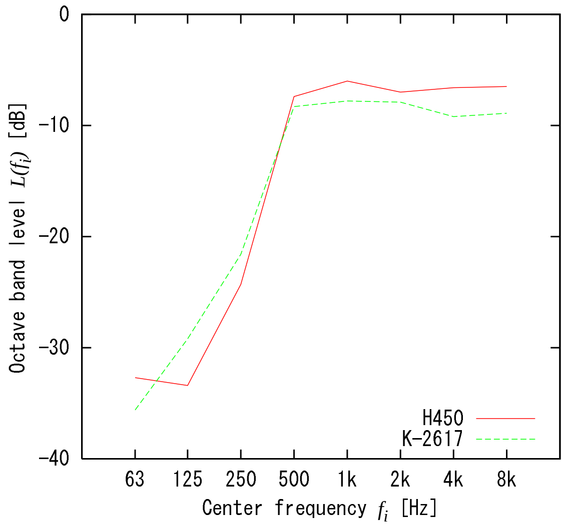
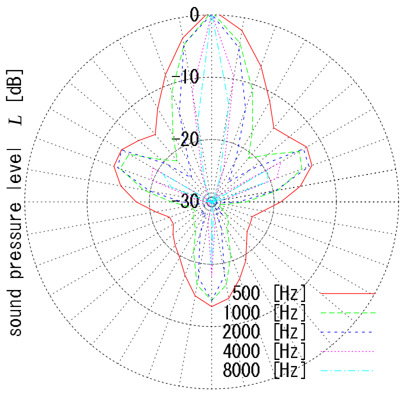
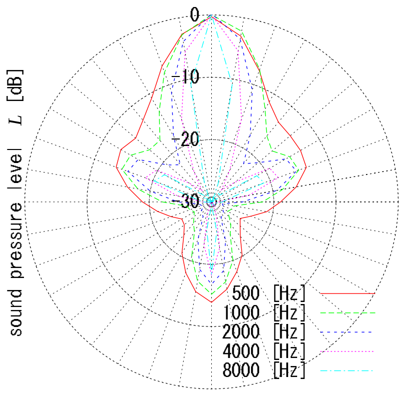


Figure 2 – Magnitude frequency response of parametric acoustic array speaker



(A) H450



(B) K-2617

Figure 3 – Directivity patterns for parametric acoustic array speaker

2.6 Result

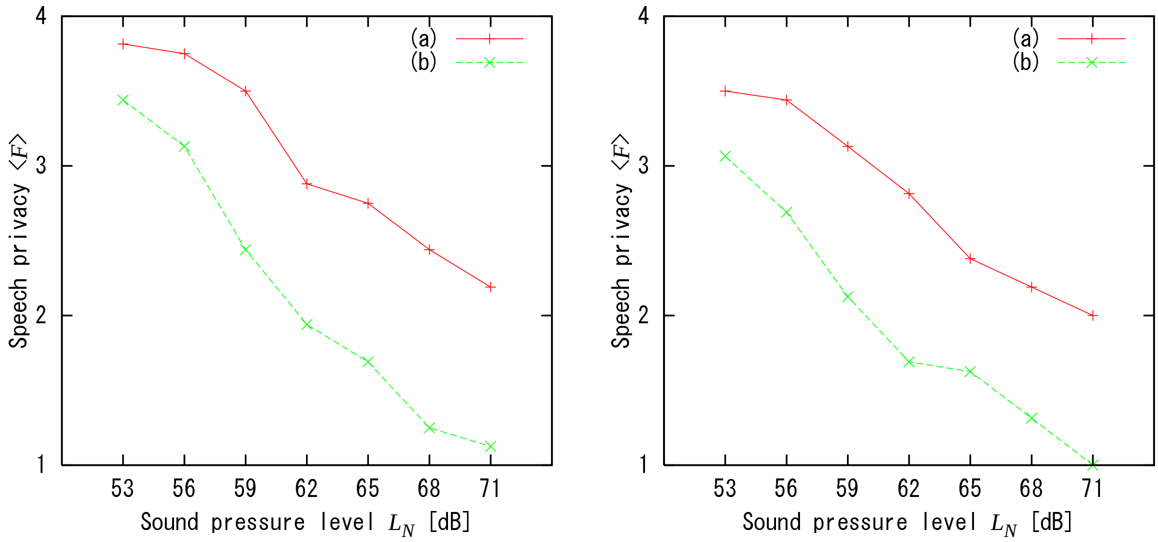
Figure 4 (A) and (B) show the relationships between sound pressure level of masking noise which transmitted from frontal directions and the speech privacy evaluation for male and female speech. Points in these figures represent averaged values of the evaluations. Figure 5 (A) and (B) show the relationship between sound pressure level of masking noise which transmitted from frontal directions and the speech privacy evaluation for male and female speech. These figures reveal that the speech privacy evaluation of parametric speaker array is smaller than that of cone loudspeaker due to the narrow frequency response at the same sound pressure level of masking noise.

3. EVALUATION INDEX TO SET UP THE DEGREE OF SPEECH PRIVACY

In this paper, an evaluation index, S [9], was introduced for setting up degree of speech privacy. S was calculated as follow:

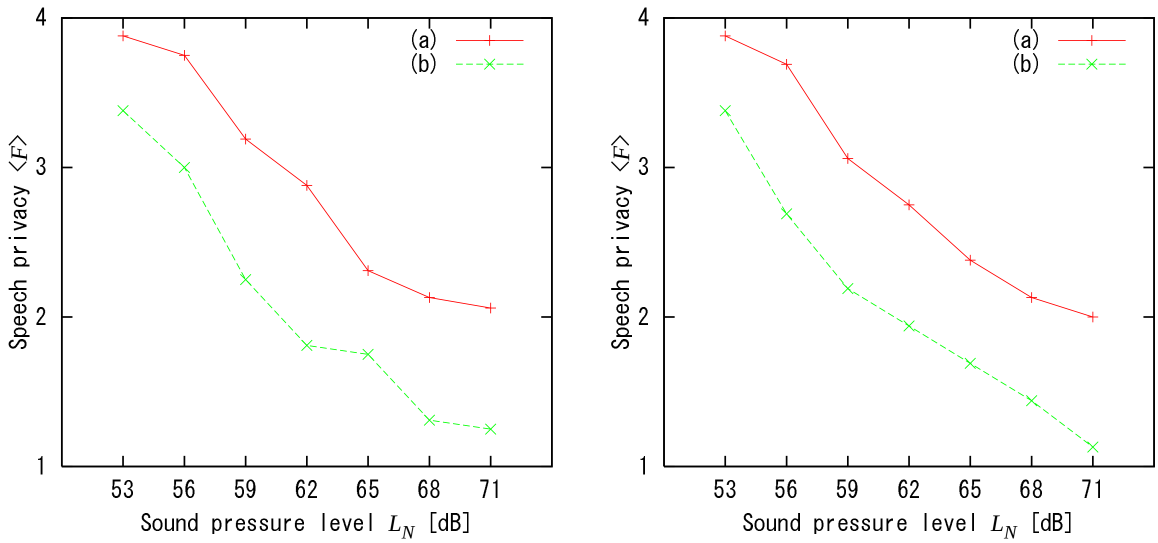
$$S = \sum_{i=1}^{8} a_i \{L_S(f_i) - L_N(f_i)\} \tag{1}$$

where a_i denotes the weight considered to percentage of 20 frequency bands [6] that contribute equally to speech intelligibility. These are included in octave bands with center frequency f_i ($f_1 = 63, f_2 = 125, \cdots, f_8 =$



(A) Male speech (B) Female speech

Figure 4 – Relationships between sound pressure level of masking noise which transmitted from frontal directions and the speech privacy evaluation



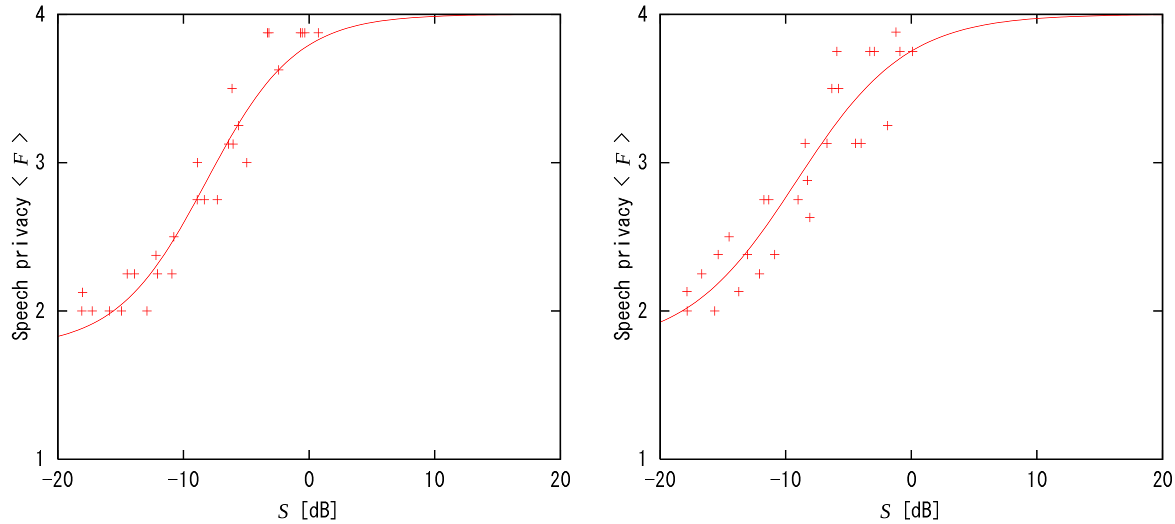
(A) Male speech (B) Female speech

Figure 5 – Relationships between sound pressure level of masking noise which transmitted from above directions and the speech privacy evaluation

8000 Hz) and are shown as follows:

$$\begin{aligned} &a_1 = 0.00, \;\; a_2 = 0.00, \;\; a_3 = 0.06, \;\; a_4 = 0.14 \\ &a_5 = 0.23, \;\; a_6 = 0.32, \;\; a_7 = 0.23, \;\; a_8 = 0.22 \end{aligned} \tag{2}$$

$L_S(f_i)$ denotes the band level with center frequency f_i ($f_1 = 63, f_2 = 125, \cdots, f_8 = 8000$ Hz) of the speech peaks. In this paper, the maximum band levels of the speech, measured by a sound level meter along with a real-time octaveband analyzer with FAST dynamic response for 30 s, are adopted as the band levels of the speech peaks. $L_N(f_i)$ denotes the band level with center frequency f_i ($f_1 = 63, f_2 = 125, \cdots, f_8 = 8000$ Hz) of the masking noise. These band levels of the noise were measured by a sound level meter along with a real-time octave-band analyzer with FAST dynamic response. We considered what was an effective index for deciding the sound pressure level of masking noise necessary to achieve speech privacy.



(A) Frontal direction (B) Above direction

Figure 6 – Relationship between S and the speech privacy evaluation in the case of transmitting masking noise from the frontal and above directions

3.1 Relationships between the index and evaluations of speech privacy

Based on the objected data from psychological experiment I, the relationships between an index and the averaged value of the evaluations for every noise condition were investigated. Figure 6 (A) and (B) show the relationship between S and the speech privacy evaluation. in the case of transmitting masking noise from the frontal and above directions. Points in this figure represent averaged values of the evaluations. Dotted line represents regression function selected by Akaike information criterion (AIC) [10] from the following types of model describing regression between them:
Logistic function:

$$y = \frac{k-c}{1+a\exp^{-bx}} + c \tag{3}$$

Modified exponential function:

$$y = (k-c)(1-\exp^{-\frac{x-b}{a}}) + c \tag{4}$$

Gonperz function:

$$y = (k-c)\exp^{-\exp^{-a(x-b)}} + c \tag{5}$$

As the index S is used, it becomes easy to set up the degree of speech privacy of 2, 3, or 4. Figure 4 also shows that when S is almost -15 dB, evaluation of speech privacy was 2, which means that it was possible to recognize the sound as a voice, though the contents of the speech were not understood.

4. OUTLINE OF PSYCHOPHYSIOLOGICAL EXPERIMENT II

To examine what psychological evaluation of speech privacy was obtained when a masking noise wastransmitted with angled parametric array speaker, Experiment II was conducted. In outline, Experiment II was as follows.

4.1 Subjects

The subjects were 14 male and 2 female students with normal hearing, and were different from the subjects in Experiment I.

4.2 Audio signal

The same audio signal as presented in Experiment I.

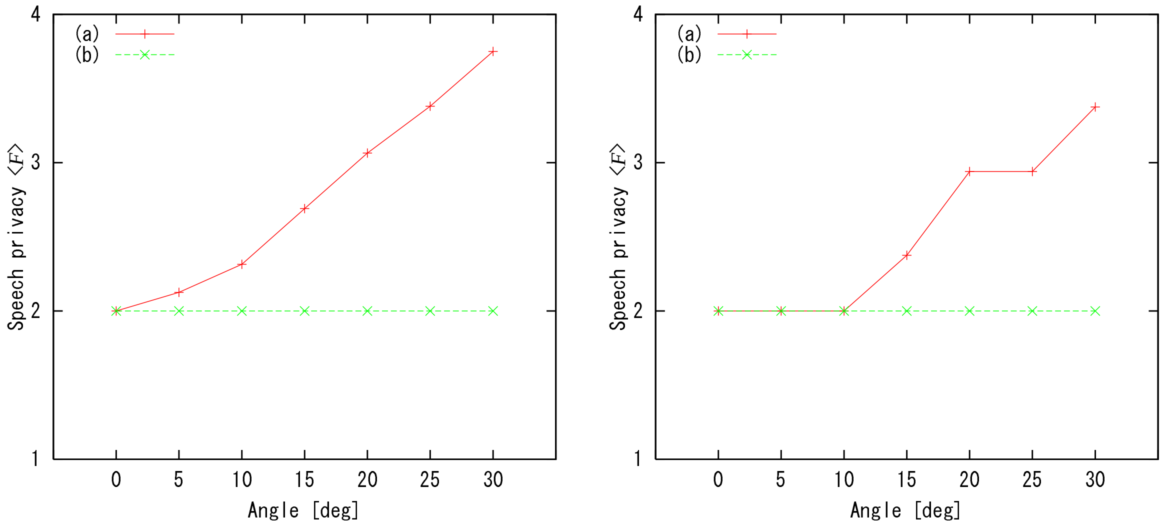


Figure 7 – Relationship between presentation angle of masking noise and the speech privacy evaluation

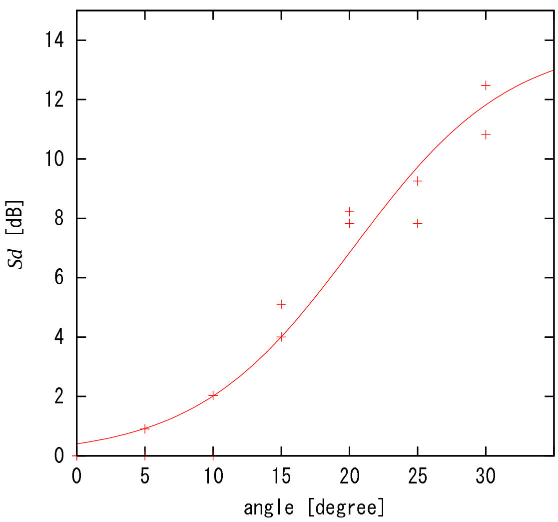


Figure 8 – Relationship between presentation angle of parametric speaker array and S_d

4.3 Masking noise

The same masking noise was used as Experiment I. From the relationship between S and evaluation (Fig. 6), the sound pressure level of the masking noise was set up so that the spectral distance could be set to -8, -15 dB which corresponded to the cases of evaluations of speech privacy of 3 and 2.

4.4 Acoustic loudspeakers

The same speakers were used as Experiment I.

4.5 Measurement

Although the same method was used as in Experiment I, a psychological impression for speech privacy was investigated when the parametric array speaker turned from frontal direction to 30 degree (right or left) at 5 degree interval.

4.6 Result

Figure 7 (A) and (B) show the relationship between presentation angle of parametric speaker array and the speech privacy evaluation for male and female speech. In these figures, as presentation angle of a parametric speaker array is larger, the speech privacy evaluation is larger. The results suggested that it is possible to maintain speech privacy in the narrow area by presenting highly directional masking sound.

Furthermore, we tried to convert degree of effect on the speech privacy evaluation of the presentation angle of parametric speaker array into value of S_d, based on the S in the case of 0 degree. Figure 8 shows relationship between S_d and presentation angle of parametric speaker array. There is a tendency for the value of S_d to increase as the presentation angle is larger. It was observed in the case of 30 degree the value of S_d is almost 12 dB, and, that effect on the speech privacy evaluation of masking noise is the smallest.

5. CONCLUSION

This study considered the highly directional sound from modulated ultrasound as a masking noise for achieving speech privacy in the narrow area. Psychological experiments were conducted in which the masking sound was transmitted to participants from frontal or above directions with a parametric acoustic array speaker. Using the experimental data, the relationships between the degree of speech privacy and the index S were investigated for setting up degree of speech privacy evaluation. The results suggested that it is possible to maintain speech privacy in the narrow area by presenting highly directional masking sound.

ACKNOWLEDGMENTS

This study was supported in part by the Japan Society for the Promotion of Science for Young Scientists (B), No.21700238, 2009 and the Japan Society for the Promotion of Science for Young Scientists (B), No. 24700208, 2012.

REFERENCES

1. W. J. Cavanauch, W. R. Farrell, P. W. Hirtle, B. G. Watters, Speech Privacy in Buildings, The Journal of the Acoustical Society of America, 34(4), 475-492, 1962.

2. A.C.C.Warnock, Studies of Acoustical Parameters in Open-Plan Offices, The Journal of the Acoustical Society of America, 63(3), 832-840, 1978.

3. J.A.Johnson, A Simplified Articulation Index Calculation for Open-Plan Spaces, The Journal of Sound and Vibration, 14-17, 1980.

4. R.Farrell, Masking Noise Systems in Open and Closed Spaces, Journal of the Audio Engineering Society, 19(3), 197-201, 1971.

5. T.Tamesue, S.Yamaguchi, T.Saeki, Study on achieving speech privacy using masking noise, The Journal of Sound and Vibration, 297(3-5), 1088-1096, 2005.

6. P. J. Westervelt, Parametric acoustic array, The Journal of the Acoustical Society of America 35(4), 1063, 1963.

7. T. Saeki, T. Tamesue, S. Yamaguchi, K. Sunada, Selection of meaningless noise for masking of speech, Applied Acoustics 65(29), 203?210, 2004.

8. Y. Tokita, Sound Environment and Control Technology II, Fuji Techno-System Co., Tokyo, 1999 (in Japanese).

9. 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) , 519?533, 2004.

10. H. Akaike, A new look at the statistical model identification, IEEE Transaction on Automatic Control AC-19 (6), 716?723, 1974.