



**Acoustics'08
Paris**
June 29-July 4, 2008

www.acoustics08-paris.org

Experimental study on applicability of sound masking system in medical examination room

Kanako Ueno^a, Hyojin Lee^a, Shinichi Sakamoto^a, Atsuko Ito^b, Mai Fujiwara^b
and Yasushi Shimizu^b

^aInstitute of Industrial Science, University of Tokyo, Komaba 4-6-1, Meguro-ku, 153-8505
Tokyo, Japan

^bCenter for Advanced Sound Technologies, Yamaha, 203 Matsunokijima, 4380192 Iwata,
Shizuoka, Japan
ueno@iis.u-tokyo.ac.jp

Recently, speech privacy to avoid oral information leakage in healthcare facilities has become an important issue. This study investigated effectiveness of sound masking system in regard to masking efficiency and annoyance for medical consultation rooms in an experimental approach. Considering actual application, two adjacent medical consultation rooms partitioned by a low sound insulation wall in a typical healthcare facility were selected as an experimental field and sound masking system was temporally installed. In the rooms, acoustic environment was measured and reproduced in an anechoic room with a 3-D sound field simulation system using a 6-ch sound recording/reproduction technique. In the simulated acoustic condition, subjective tests were designed to quantify the masking efficiency and annoyance caused by the masking sound. The annoyance test was conducted in listening condition (with high attention to the sound) and in talking condition (with low attention). As a result, mixed maskers composed by water stream, synthesized speech signals, and steady state noise showed high performance in both aspect of masking efficiency and annoyance.

1 Introduction

In the medical consultation room, most people would not like anybody outside listen to their private conversation. At the same time, they would not feel comfortable to hear the other's personal conversation on health. This issue, speech privacy in healthcare facilities, has been already discussed in the interim design guideline of healthcare facility in the US [1]. In Japan, the issue has not been regarded as important in general. However, among medical staffs, the necessity of "confidential" in the speech privacy has becoming concerned. In our inquiry survey in a medical facility, some of the doctors and nurses commented that the patient's personal information in their consulting conversation should be kept in secret. As a practical action, they actually made the voice softer to avoid the conversation leaking to the next room when they noticed somebody enter the adjacent room. This example indicates that oral information leakage between consultation rooms and to waiting lobby has been becoming a serious issue.

As discussed in the design guideline in the US, appropriate sound isolation is primarily important to avoid leakage of oral information. However, considering the actual medical activity of doctors and nurses, spatial isolation by doors with high sound isolation performance would not be accepted in some of the healthcare facilities. In the case, the use of sound masking system would be beneficial if it provides the speech privacy and spatial continuity for medical activities at the same time.

This study investigates a practical solution using sound masking system to secure the speech privacy in Japanese healthcare facility. A typical medical consultation rooms with low sound isolation is selected for an experimental field and a sound masking system is temporally installed. The acoustical environment is measured and reproduced for listening tests that examine the masking effect and annoyance produced by the masking sound. This paper presents the results of the listening tests with different types of masking sound and discusses the method to produce high performance maskers.

2 Experimental method

2.1 Sound field model

In order to investigate the effect of the masking sound in room environments, we assume three components—target,

masker, and background noise (see Figure 1). The target is the speech signal to be masked and it transmits to the listener with the sound transmission property denoted by IR_T . Similarly, the masking sound reproduced by a loudspeaker (masker) transmits to the listener with the sound transmission property IR_M . Listeners also hear the background noise of the surrounding environment. We assume that the experimental sound fields comprise a combination of these three components.

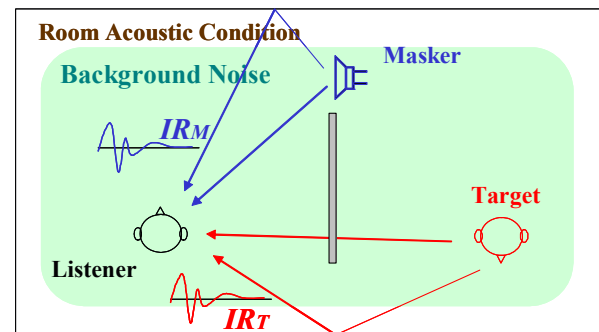


Fig.1 Components of sound field for the experiment.

2.2 Tested room

As a typical site, we focused on an actual medical consultation rooms where the sound insulation between rooms is insufficient to secure the speech privacy. Figure 2 shows the dimension of spaces and settings of sound sources and a receiver (listener). A listener is set in a room and the target is placed in the adjacent room. The masker is reproduced from four loudspeakers set at the corners.

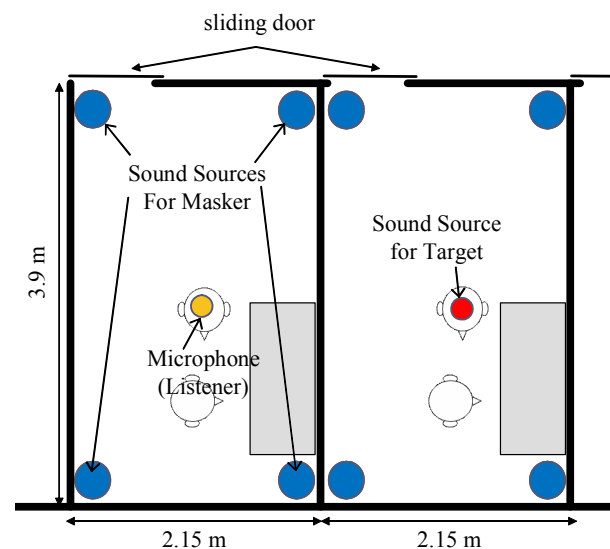


Fig.2 Dimension of spaces and measurement positions.

The sound insulation property between two rooms is shown in Fig.3. In the room, conversation in the adjacent room was easily heard and understood.

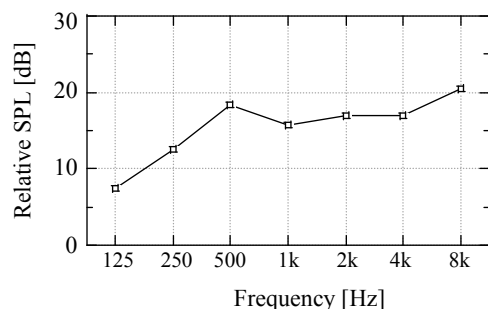


Fig.3 Sound insulation property between rooms under examination.

2.3 Sound field simulation

A 6-ch sound field simulation system was employed to simulate actual sound environments in an anechoic room with 3-D sound information. In this system, sound fields are recorded through a 6-ch microphone system (Sanken CU6ch) comprising six unidirectional microphones with cardioid directional characteristics, which are combined at every 90° in the horizontal and vertical planes. The 6-ch signals are reproduced through six loudspeakers (TANNOY, T12) arranged on a spherical surface with a radius of 2 m in the anechoic room (7 meters on each side). One experiences a natural 3-D auditory sensation when hearing the reproduced sounds at the center position. [2]

In order to obtain the acoustic data of the actual room environment, acoustical measurements of the impulse responses and background noise were conducted in the medical consultation rooms. The positions of sound sources and microphones are shown in Fig. 2. For the masker, four dodecahedral omnidirectional loudspeaker systems (Solid Acoustics SA-355sn) are attached to the ceiling at the four corners in each room. For the target, a dodecahedral omnidirectional loudspeaker system (TS-12M) is set at a height of 1.2 m in the adjacent room. For the receiver, the 6-ch microphone system and a omnidirectional microphone to measure the acoustical characteristics are set at a height of 1.2 m in the room. The impulse responses are measured by the swept-sine technique. As a background noise (BGN), air conditioning system noise is recorded by the 6-ch microphone system at the listener's position.

Figure 4 shows the diagram of the sound reproduction system for three sets of the 6-ch sound signals for the target, masker, and background noise. To provide the transmission properties of the target and masker, the 6-ch directional impulse responses, IR_T and IR_M , measured by the 6-ch microphone system are convolved using a 6-ch real-time convolution system. The background noise is reproduced by a hard disk recorder. The three components are mixed for each channel (direction) and provided to the loudspeakers. In the listening tests, a subject is seated at the center position of the experimental sound field and listens to the test sound field.

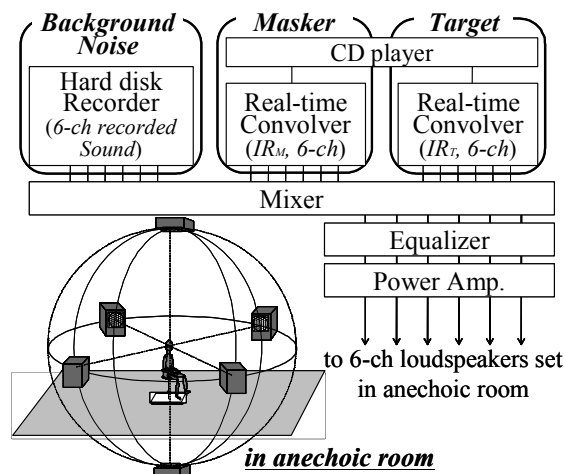


Fig.4 Diagram of 6-ch sound field reproduction system.

2.4 Test sound field

Test sound field was composed by the combination of sound components of the target, masker, and background sound. The sound level to present each experimental condition was adjusted by measuring the A-weighted SPL at the center position of the simulated sound field by each component.

The target sounds was provided by convolving a dry source (speech signal) with the 6-ch directional IRs (IR_T) and presented at 45 dB in A-weighted SPL. The level was determined considering the sound insulation performance measured at the tested room.

In our previous study, it was observed that speech-like maskers processed from speech signal had higher masking efficiency than steady-state noise maskers [3]. However, it was indicated that the speech-like maskers tended to cause higher annoyance when it was outstanding from background sound [4]. Therefore, in the current study, we examine the effect of mixture of different types of sound by intending to maximize masking effect and minimize annoyance caused by a masker.

As maskers, three sound elements were used; recorded HVAC noise (HVAC), recorded water stream (WS), and processed speech (SP). In the test, masking efficiency of each sound element of the masker and that of the mixture of two or three sound elements were measured (see Table 1). The level balance of the multiple elements of the mixed maskers was decided by the on-site experimenters so that the mixed sound had higher masking effect and lower annoyance. Besides, two types of steady state noise masker having the same frequency characteristics of WS and SP (ID 4 and 5 in Table 1) were additionally examined (denoted by WS-N and SP-N, respectively). Each masker was convolved with the measured 6-ch IRs (IR_M). SPLs of each masker were changed in 5 dB steps from 35 to 55 dB (-10 to 10 dB in T/M ratio) except that the HVAC noise, whose level was changed from 40 to 60 dB (-15 to 5 dB in T/M ratio).

Figure 5 shows the frequency characteristics of each sound element measured at the center position of the experimental sound field, where SPL of each element was set for ID 10 masker.

As a background noise (BGN), the air conditioning system noise recorded by the 6-ch microphone system was reproduced at 41 dB.

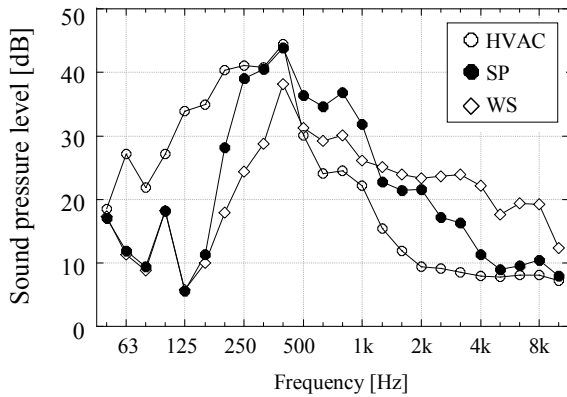


Fig. 5: Frequency characteristics of each sound element of the masker

3 Exp.1: Word intelligibility test

3.1 Procedure

Masking efficiencies of maskers were measured by means of a word intelligibility test in varying the SPL ratios (T/M ratios) of its target to masker in dB.

For the target speech source, the set of test words were chosen from “NTT Database Series: Lexical Properties of Japanese” [5]. Out of the database, five mora (phonetic word length of Japanese) words with the same word familiarity were chosen as target words. Eight words recorded with four persons (two words each, three female and one male) are presented in each experimental condition.

Target words were presented one by one with a four-second blank to write down the word they heard. The subjects are encouraged to guess the word. Eight subjects with normal hearing ability participated in the experiment.

3.2 Results

The condition mean of percent correct responses of the word intelligibility test across the subjects was calculated for each target/masker condition (8 words * 8 subjects=64 responses). Furthermore, the relationship between the T/M ratio and the word intelligibility score was modeled for each of the maskers by using logistic regression analysis. As examples of the results, Fig. 6 shows the condition means of percent correct responses with dots and the logistic regression model with lines.

Based on the logistic model for each masker, the predicted value of the T/M ratio of 20 % word intelligibility (80 % incorrect), which is supposed to be the case where the masker considerably masks the target, is calculated. The results for each masker are shown in Table 1.

From the results shown in Table 1, followings are indicated.

- Comparing three sound elements (HVAC, WS, and SP), WS has the highest masking efficiency, whereas HVAC has the lowest. The difference of SPL between

the highest (WS) and the lowest (HVAC) having the same word intelligibility score (20 %) is about 16 dB.

- Comparing the sound elements (WS and SP) and steady state noise (WS-N and SP-N), the efficiency of steady state noises decreases and about 6dB higher SPL from the original signal is required to be the same value in masking efficiency. This indicates that the fluctuation in time domain of WS and SP has considerable contribution to the masking efficiency.
- Comparing two steady state noise maskers, WS-N has 5 dB higher efficiency than SP-N. This fact indicates that high frequency component (2k to 8k Hz octave bands) may have higher masking effect.
- With regard to the effect of mixture of elements, the mixed maskers having relatively higher performance show comparable efficiency with WS.

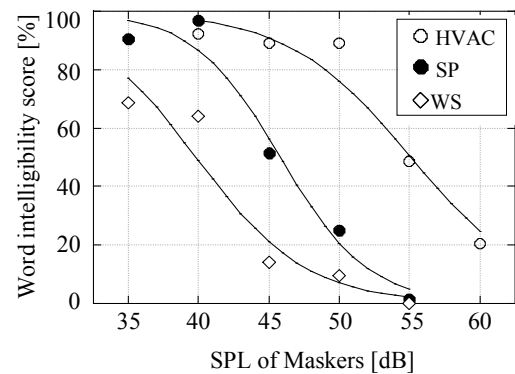


Fig. 6: Results of the word intelligibility test with each single sound element of the masker (HVAC, WS, and SP).

Table 1 Predicted SPL (A-weighted, [dB]) of maskers and background noise (BGN) at 20 % word intelligibility

ID	Components of masker	SPL of maskers			SPL of BGN	SPL in total
		HVAC	WS	SP		
1	One element	-	45.3	-	-	45.3
2		-	-	50.1	-	50.1
3		61.2	-	-	-	61.2
4		-	51.8* ¹	-	-	51.8
5		-	-	56.6* ²	-	56.6
6	Two elements	-	38.3	42.7	-	47.3
7		44.0	40.3	-	-	49.9
8		45.6	-	46.7	-	51.7
9	Three elements	42.6	39.1	43.6	-	48.3
10		42.0	38.5	43.0	41.0	47.6

*1: steady state noise with frequency characteristics of WS (WS-N)

*2: steady state noise with frequency characteristics of SP (SP-N)

4 Exp.2: Subjective judgment test

4.1 Procedure

In order to quantify the annoyance caused by the complex sound that comprises the target, masker, and background noises, a subjective judgment test was designed. In addition,

subjective impression on masking effect is also questioned. Table 2 shows outline of the questionnaire.

Table 2 Questionnaire asked in the subjective test

Q1.*1 Did you concentrate on the conversation? (5 steps)
Q2. Was the room noisy? (5 steps)
Q3. Was the sound environment preferable as a medical consultation room? (5 steps)
Q4. Did you hear any speech from surroundings? (Yes/No) if Yes,
Q4a. Did you understand the meaning? (5 steps)
Q4b. Was the speech annoying? (4 steps)
Q5. If you had any annoying sound except speech, check the sound (multiple choices), answer what extent you wanted to get it off (5 steps), and what extent it was annoying (3 steps).

*1: asked only in talking condition

The subjective judgment tests were conducted in two situations; listening condition (as high attention to the sound) and talking condition (as low attention to the sound). In the listening condition, a subject was requested to listen to the sound for 15 seconds and then start to write down one’s judgment on the questions in the test sound condition. In the talking condition, a subject was requested to have a conversation with an experimenter (female) for 60 seconds and start the judgment. In both sessions, subjects were asked to image that he/she was in a medical consultation room. Twelve subjects with normal hearing ability participated in this experiment.

The recorded conversation between two female were used as a target source signal. The maskers already examined in the experiment 1 were tested. Those were reproduced at the SPL value having 20 % word intelligibility (ID 1-3, 6-10 in Table 1) except steady state noise maskers (ID 4 and 5) . In addition, single elements (3 conditions), only BGN, mixture of two elements (3 conditions) and mixture of three elements (1 condition) are also tested by the level setting of ID 10.

4.2 Results on masking effect

Figure 7 shows the condition means of the “degree of understanding the meaning” (asked by Q4a) across subjects for the masker conditions in the relationship with the word intelligibility score. From fig.7, following results are indicated.

- The subjective judgments are considerably different between two situations; listening and talking. In the talking condition, the “degree of understanding” is generally low.
- Even if the maskers have the same word intelligibility score (20 %), a degree of subjective judgment on masking effect are considerably different (see Fig.7-a). Comparing three sound elements (HVAC, WS, and SP), SP is judged as the most effective and WS is judged as the least effective.

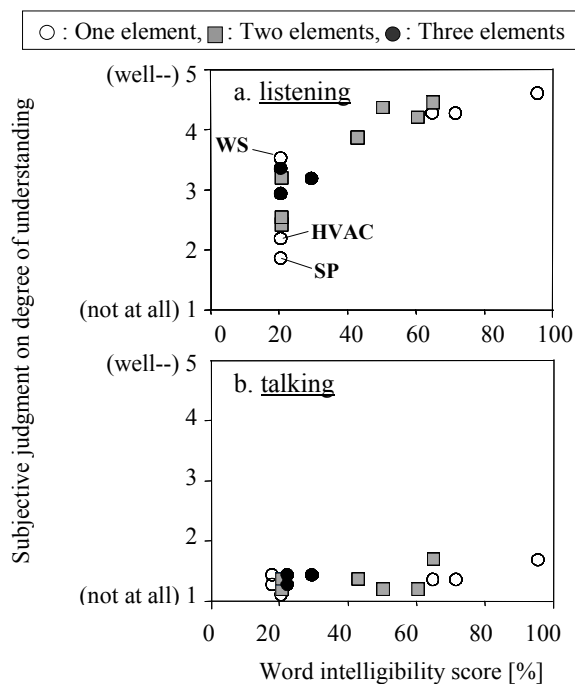


Fig. 7: Relationship between word intelligibility score and subjective judgments on masking effect

4.3 Results on annoyance

Figure 8 shows the condition means of the degree of noisiness (asked by Q2) across subjects in the relationship with SPL of maskers. Here, only the results of the maskers having 20 % word intelligibility score are plotted.

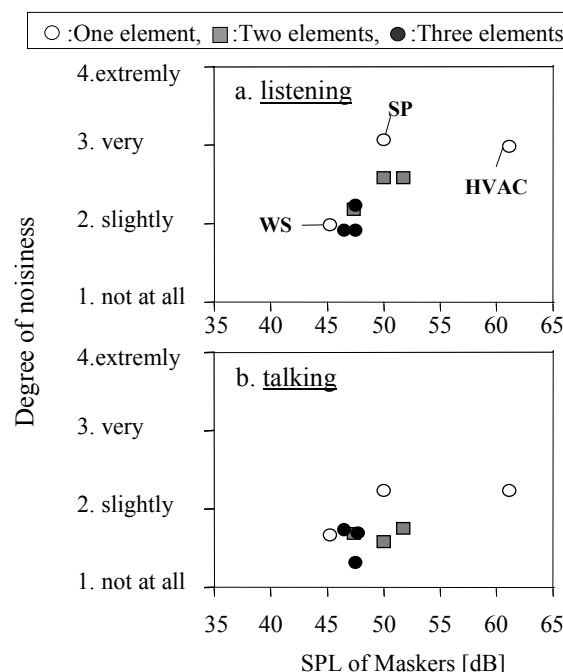


Fig.8 Relationship between SPLs of maskers and subjective judgments on noisiness

From the results shown in Fig. 8, followings are indicated.

- The subjective judgments on noisiness are considerably different between two situations;

listening and talking. In the talking condition, the presented sound environments are judged less noisy in comparison with the listening condition.

- Subjective judgment on noisiness roughly correlates to SPL. However, as shown in the results of listening condition (Fig.8-a), even if the maskers are presented in the similar SPL value (see the dots around 50 dB), subjective judgments are considerably different. SP, used as the single element, causes higher noisiness, whereas mixed maskers cause lower noisiness.

With regard to the results of the question that examines the annoyance (Q5), condition means of the degree of annoyance across subjects were calculated. If the subjects did not check the elements as an “annoying” sound, we considered the degree of “annoyance” to be 1 (not annoying at all). Figure 9 indicates the results of the degree of annoyance judged in the listening condition in the relationship with SPLs of each element, where the results of one element masker and a mixture of three elements masker with BGN (ID 10) are plotted.

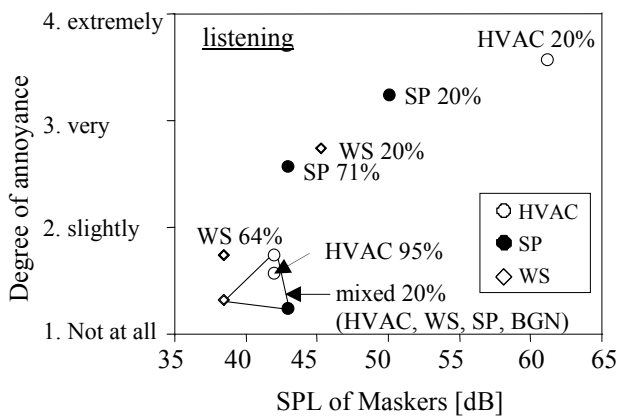


Fig. 9 Relationship between SPLs of maskers and subjective judgments on annoyance of each element.

From the results shown in Fig. 9, followings are observed.

- Mixed masker gives lower annoyance than single element maskers.
- When each element of mixed masker is reproduced separately, masking efficiency falls off (word intelligibility score become higher). Besides, annoyance of SP and WS become considerably lower when those elements are mixed. This fact indicates that a mixture of the elements has the effect on both of increasing masking efficiency and reducing annoyance.

5 Summary

Setting an experimental field of medical consultation rooms with insufficient sound insulation between rooms, acoustical environment was measured and simulated for listening tests in an experimental room and performance of sound masking system was examined.

As a result of word intelligibility test which examined the masking efficiency, it was found that water stream (WS) had the highest masking efficiency which may be caused by its peculiar fluctuation in time domain and high frequency

components. However, it was found that the water stream (WS) was given lower rating when it was evaluated by the subjective judgment test to ask the degree that the intruding conversation through walls was understandable. In the subjective judgment test, it was shown that a speech-like masker (SP) was evaluated to have high masking effect. On the other hand, in the judgment on annoyance, the SP masker was evaluated lower (higher annoyance). As a compatible solution, the mixture of signal elements was proved to be effective in both terms of masking effect and annoyance. Consequently, it was indicated that the masking system that was applicable to the medical consultation rooms could be suggested when the mix masker was designed successfully.

Another finding of the study was that the effectiveness of maskers was measured differently by the evaluation methods. That is, the results of word intelligibility test are not always in good correlation with the results of a degree of speech privacy in subjective impression. Furthermore, as a matter of course, the subjective impression was judged differently according to the degree of attention to the sound (task during the judgment). The evaluation method of masking system should be further studied by considering these facts.

Acknowledgments

The authors wish to thank the subjects who participated in this study.

References

- [1] Joint Subcommittee on Speech Privacy of the ASA/INCE/NCAC, “Interim Sound and Vibration Design Guidelines for Hospital and Healthcare Facilities,” public draft dated Nov. 28 (2007).
- [2] S. Yokoyama, K. Ueno, S. Sakamoto, and H. Tachibana, “6-channel recording/reproduction system for 3-dimensional auralization of sound fields,” *Acoustical Science and Technology* 23, 97–103 (2002).
- [3] A. Ito, A. Miki, Y. Shimizu, K. Ueno, HJ. Lee, and S. Sakamoto, “Oral information masking considering room environmental condition, Part 1: Synthesis of Maskers and examination on their masking efficiency,” *Proceedings of INTER-NOISE 2007* (2007).
- [4] K. Ueno, HJ. Lee, S. Sakamoto, A. Ito, A. Miki, and Y. Shimizu, “Oral information masking considering room environmental condition, Part 2: Subjective assessment for “Masking efficiency and Annoyance”,” *Proceedings of INTER-NOISE 2007* (2007).
- [5] S. Amano and T. Kondo, *Nihongo no goitokusei* [Lexical properties of Japanese] (Vols. 1-6, NTT database series, Tokyo, 1999)