Proceedings of the International Symposium on Room Acoustics, ISRA 2010

29–31 August 2010, Melbourne, Australia

The relationship between sound insulation performance of walls and word intelligibility scores

Yasushi Hoshino (1), Hayato Sato (2), Masayuki Morimoto (2) and Yasuhiko Odagawa (2)

(1) Environmental Acoustics Laboratory, Kobe University, Rokkodai, Nada, Kobe, JAPAN
/ Nippon Sheet Glass Environment Amenity Co,. Ltd., Takanawa, Minato, Tokyo, JAPAN
(2) Environmental Acoustics Laboratory, Kobe University, Rokkodai, Nada, Kobe, JAPAN

PACS:

ABSTRACT

Conversation is required to be shielded from someone in an adjacent room if it includes confidential information. Word intelligibility tests were performed in a total of 185 sound fields to examine the relationship between sound insulation performance and the degree of conversation leakage. The parameters of the test sound fields were background noise level in the next room and the sound pressure level difference between two rooms. The background noise level was changed from 30 to 50 dBA. The sound pressure level difference was parametrically changed in terms of frequency characteristics (8 kinds) and absolute values (10 kinds). The results showed that word intelligibility scores were strongly correlated with A-weighted speech-to-noise ratio and SNR_{uni32}. A multiple logistic regression analysis demonstrated that word intelligibility scores can be estimated with high accuracy from the weighted level difference and A-weighted background noise level.

INTRODUCTION

Conversation is one of essential forms of communication, and is frequently done everywhere in our everyday life. However, conversation sometimes includes confidential information that should be shielded from third persons. The rooms where there is possibility that people talk about something confidential, such as consulting rooms in banks or hospitals or pharmacies, meeting rooms in offices, and so on, should be designed considering leakage of confidential information by speech transmitted through boundary walls and other paths.

The terms of "speech privacy" or "speech security" are often used for the topic of quantifying the leakage of confidential speech, and several studies have examined it in detail. Cavanaugh *et al.*[1] performed privacy rating tests using the simulated speech sounds transmitted through 5 types of walls, with the additional noise corresponding to NC-35, and demonstrated how the privacy rating related to Articulation Index[2]. This indicates that the privacy rating strongly relates to intelligibility scores.

The relationship between intelligibility scores and sound insulation performance has been investigated based on listening tests similar to those by Cavabaugh *et al.* Gover *et al.*[3] suggested SNR_{uni32}, that is a frequency-weighted average signal-to-noise ratio with uniform frequency weightings, for estimating audibility, cadence, and intelligibility which relate to speech security performance. Park *et al.*[4] compared speech intelligibility scores with sound insulation performance expressed by STC (Sound Transmission Class) from the ASTM E413 standard[5] and R_W (Weighted Sound Reduction Index) from the ISO 717-1 standard[6], and demonstrated that SNR_{uni32} was more suitable than STC and R_W for estimating speech intelligibility scores.

The previous studies clearly showed that speech-to-noise ratio, in other words, background noise level and sound insulation performance, are important variables for evaluating speech privacy or security performance. Furthermore, the two variables can be controlled in acoustic design of rooms. Therefore, it is useful for speech privacy or security performance to be estimated from the two variables. However, the previous studies did not use background noise level as a listening test parameter. Cavabaugh *et al.* varied background noise level in the preliminary test, but the range was only 10 dB. Gover *et al.* varied frequency characteristics of background noise, but background noise level was constant at 45 dBA. Accordingly, estimate equations of the performance separately including background noise level as a variable have not been proposed before.

In the present study, word intelligibility tests were performed to clarify the relationship among sound insulation performance, background noise level, and the degree of conversation leakage. The parameters of the tests were background noise level and the sound pressure level difference between two rooms. The effect of room acoustics is not considered in the present study to simplify listening tests, based on the results by Bradley *et al.*[7] that indicated results without room acoustics will be on the safe side from the viewpoint of evaluation of conversation leakage.

METHODS

Situation

Figure 1 illustrates the situation assumed in the present study. The people A and B are talking about something confidential. Person C is not an eavesdropper, but can hear speech from the next room if sound insulation performance between the two rooms is not enough high. It is assumed that person A is speaking with "Normal" vocal effort[8], and speech level at the position of person B is 58 dBA. The sound pressure level difference (D) between the positions of B and C was used as the parameter of sound insulation performance. Equation

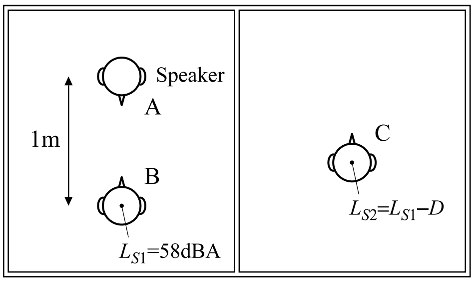

Figure 1: The situation assumed in the present study.

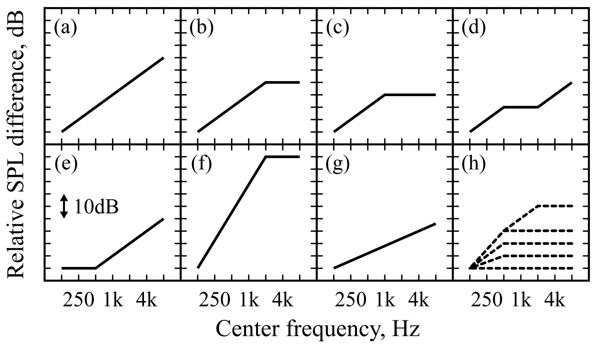

Figure 2: The frequency characteristics of the sound pressure level difference used in the present study.

1 represents the definition of D in the present study. L_{S1} and L_{S2} are speech level at the positions of the people B and C, respectively.

$$D = L_{S1} - L_{S2} \quad (1)$$

Sound pressure level difference (D)

Frequency characteristics and absolute values of D were used as parameters of listening tests. Figure 2 represents the frequency characteristics of D used in the present study. The frequency characteristics from (a) to (g) were representatives of transmission loss of walls. The characteristics were determined based on sound insulation characteristics of walls modeled by Tachibana *et al*[9], and also based on sound insulation data from the text book by Maekawa and Lord[10]. The frequency characteristic of (h) is the curves used in Japan to determine the sound insulation rank[11], and changes depending on its rank (see Fig. 3).

Absolute values of D with each frequency characteristic were varied to satisfy 10 kinds of the sound insulation rank from D_r-0 to D_r-45. Figure 3 represents the reference curve for the sound insulation rank used in Japan. The curves for D_r-0, D_r-5 and D_r-10 are not included in the original reference curve chart, and are defined in the present study for convenience. The dashed line in Fig. 3 represents D with the frequency characteristic of (a) which satisfies the rank of D_r-25. In order to satisfy the rank of D_r-25, D in all frequency bands must exceed the reference curve of D_r-25. Eighty kinds of D, which were the combinations of 8 kinds of the frequency characteristic and 10 kinds of the sound insulation rank, were used in the present study.

Speech stimuli

A total of 148 Japanese words were used as test words. The test words were selected from the familiarity-controlled word lists by Sakamoto *et al*.[12] to be most familiar to both young and elderly people. The test words consisted of four syllables, and were spoken by a female Japanese in an anechoic room.

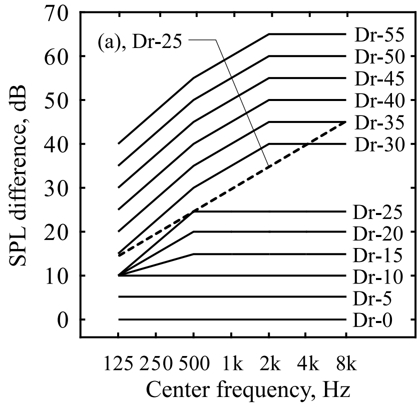

Figure 3: The reference curve of the sound insulation rank[11]. For example, the dashed line represents the sound pressure level difference with the frequency characteristics of (a) which satisfies D_r-25. The ranks from D_r-0 to D_r-10 are not included in the original chart.

Sakamoto *et al*. reported that word intelligibility scores increased with increasing word familiarity. Therefore, from the view point of evaluation of confidential information leakages, using the most familiar words provides evaluation on the safe side. Each test word was filtered to reduce its octave band level according to 80 kinds of D described above, and to make speech stimuli that simulate speech sounds at the position of person C in Fig. 1.

Listening tests were performed in an anechoic room. Figure 4 represents the loudspeaker arrangement used in the listening tests. The speech stimuli were presented from the loudspeaker in front of the listener at a distance of 1.5 m. Table 1 represents the presentation level of the speech stimuli, i.e. L_{S2} for each D. The presentation levels were measured using a sound level meter at the position of the center of the listener's head, in the absence of the listener. $L_{Amax,slow}$ for each speech stimulus was measured while the stimulus was played repeatedly, and was set at the levels shown in Table 1.

The shaded D values in Table 1 indicate that a particular combination of the absolute value and the frequency characteristic satisfies more than one rank at the same time. For example, D for the rank of D_r-10 with the frequency characteristics of (a) also satisfies the ranks of D_r-15 and D_r-20. Therefore, only D with the highest rank of the shaded D for each frequency characteristic was used in the listening tests. In other words, 67 of the 80 kinds of D were actually used.

Background noise

To simulate general room noise, a steady-state random noise with -5 dB per octave decay in frequency domain was added to each speech stimulus. The additional noise was presented from the five loudspeakers shown in Fig. 4 at the same time. Five different noise signals, which were uncorrelated with each other but had the same frequency characteristic, were presented from the respective loudspeakers in order to make the degree of inter-aural cross correlation of the additional noise close to the theoretical value of the correlation coefficient between two different points (distance: 0.3 m) in a diffuse sound field[13]. The L_{Aeq} of the additional noise was measured at the same position as that of the speech stimuli, and was set in five steps of 5 dB from 30 to 50 dBA. The additional noise of X dBA is abbreviated as NX in the rest of this paper. For example, N40 represents the additional noise of 40 dBA.

The combinations of D and the additional noise level with A-

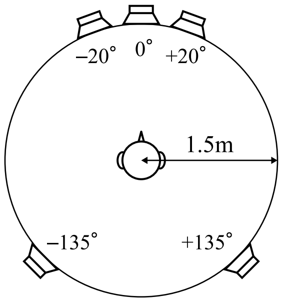

Figure 4: Loudspeaker arrangement.

Table 1: Presentation level of speech stimuli for each sound pressure level difference.

Rank	Frequency characteristic							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
D_r-0	47.4	47.4	47.8	49.0	54.9	41.0	50.8	58.0
D_r-5	42.4	42.4	42.8	44.0	49.9	36.0	45.8	53.0
D_r-10	37.4	37.4	37.8	39.0	44.9	31.0	40.8	48.0
D_r-15	37.4	37.4	37.8	39.0	39.9	31.0	40.8	43.6
D_r-20	37.4	37.4	37.8	39.0	34.9	31.0	36.8	39.4
D_r-25	32.4	32.4	32.8	34.0	29.9	31.0	31.8	35.6
D_r-30	27.4	27.4	22.8	19.0	24.9	26.0	22.8	29.2
D_r-35	22.4	22.4	17.8	14.0	19.9	21.0	17.8	24.2
D_r-40	17.4	17.4	12.8	9.0	14.9	16.0	12.8	19.2
D_r-45	12.4	12.4	7.8	4.0	9.9	11.0	7.8	14.2

(dBA)

weighted speech-to-noise ratio ($SNR(A)$) from around -20 to 0 dB were used in the listening tests. Specifically, the combinations were N50 and D with the ranks from D_r-0 to D_r-25 (36 conditions), N45 and D_r-5 to D_r-30 (36 conditions), N40 and D_r-15 to D_r-35 (34 conditions), N35 and D_r-20 to D_r-40 (39 conditions), and N30 and D_r-25 to D_r-45 (40 conditions).

Procedure

Two listening tests were performed in the presented study. Test I was for the conditions of N30 and N40 (74 conditions), and Test II was for the conditions of N35, N45, and N50 (111 conditions). Each listener was asked to write down speech stimuli as they listened using katakana characters (Japanese phonograms).

Thirty-seven listeners participated in Tests I & II, respectively. The listeners were young-adults, and had normal hearing level.

In Test I, each listener listened to 296 speech stimuli that included each test word twice, and each condition 4 times. The combination of the test words and the conditions was different for different listeners. Test I was divided into 8 sessions to listen to 37 speech stimuli. A total of 148 speech stimuli (37 listeners $\times$ 4 times) were presented in each condition after finishing Test I.

In Test II, 111 of 148 test words were used. Each listener listened to 333 speech stimuli which included each test word and each condition thrice. Test II was divided into 9 sessions to listen to 37 speech stimuli. A total of 111 speech stimuli (37 listeners $\times$ 3 times) were presented in each condition after finishing Test II.

RESULTS AND DISCUSSIONS

The word intelligibility score, which is the percentage of the speech stimuli written down correctly, was calculated from the results for all listeners in Tests I and II.

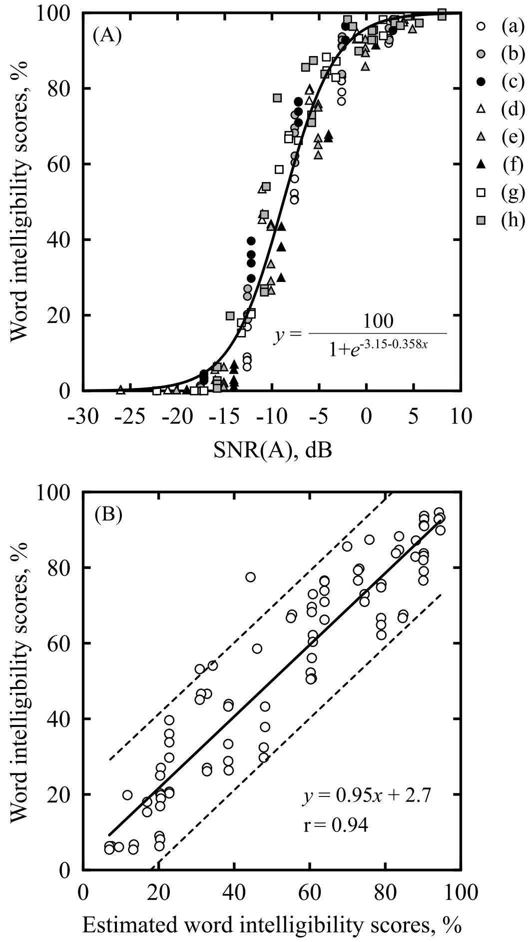

Figure 5: The relationship between word intelligibility scores and A-weighted speech-to-noise ratio ($SNR(A)$). The panel (A) represents the scores as a function of $SNR(A)$. The solid curve represents a logistic regression curve, and the different symbols represent the different frequency characteristics of sound pressure level difference between two rooms. The panel (B) represents the relationship between the scores and the estimated scores from the logistic regression curve shown in the panel (A). The solid line represents a linear regression line between the two scores, and the dashed lines represent 95% prediction intervals.

Single number evaluation

$SNR(A)$ and SNR_{uni32}[3] were calculated to compare with word intelligibility scores. $SNR(A)$ is the presentation level of the speech stimuli minus that of the additional noise level.

Figures 5 represents the relationship between the word intelligibility scores and $SNR(A)$. The panel (A) represents the scores as a function of $SNR(A)$. Different symbols represent different frequency characteristics of D shown in Fig. 2. The scores began to depart from 0% when $SNR(A)$ exceeded around -15 dB, and then increased with increasing $SNR(A)$ up to $\pm$0 dB. The scores seemed to fit in a logistic regression curve, regardless of the frequency characteristics of D. The panel (B) represents the relationship between the scores and those estimated from the regression curve. Only the scores in the range from 5 to 95% were used in a linear regression analysis. The two scores are highly correlated with each other (r=-0.94). The dashed lines represent 95% prediction intervals. The intervals were $\pm$20%, and actually the scores varied from 0 to 40% for the estimated score of around 20%.

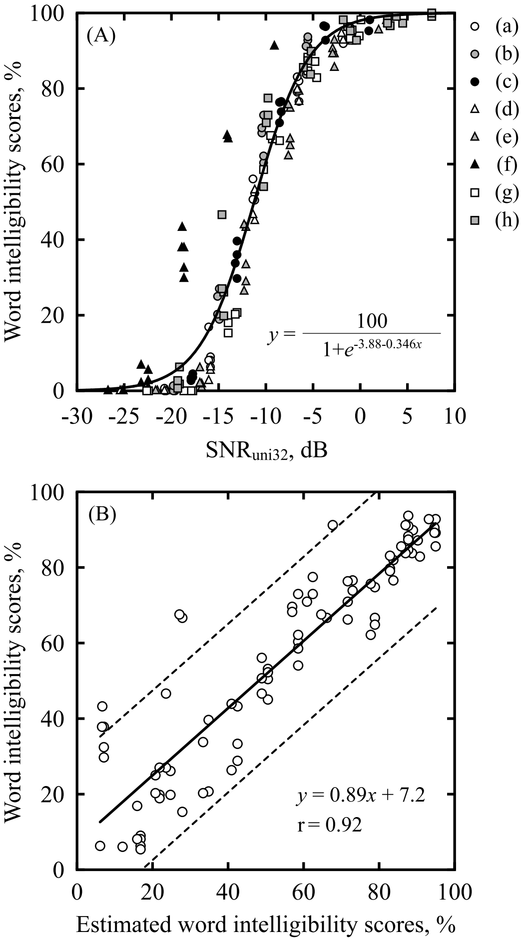

Figure 6: As for Fig. 5, but for SNR_{uni32}.

Figure 6 represents the relationship between the scores and SNR_{uni32}. The relationship was very similar to that between the scores and $SNR(A)$, except that the scores for the frequency characteristic of (f) were higher than those for other frequency characteristics. In other words, SNR_{uni32} underestimates the scores for the frequency characteristic of (f). This result suggested that the scores in the present study were not predominantly affected by the decrease of the speech-to-noise ratio at high frequencies. The frequency characteristic of (f) was steeper than that for the other characteristics, and therefore, the speech-to-noise ratio at high frequency more rapidly decreased. SNR_{uni32} can take count of the decrease at high frequencies while $SNR(A)$ cannot, because $SNR(A)$ for (f) is mainly determined by mid-frequency components. However, the scores for (f) were not affected so much by the decrease of the speech-to-noise ratio at high frequencies, and as a result, the scores for (f) did not fit in the regression curve for SNR_{uni32}, while they fit in that for $SNR(A)$.

The regression analysis without the frequency characteristic of (f) showed that the correlation coefficient between the scores and those estimated from SNR_{uni32} was -0.97, and it was higher than that for $SNR(A)$ (r=-0.95). Therefore, it is concluded that the prediction accuracy for $SNR(A)$ and SNR_{uni32} is not significantly different from each other in most cases. However, $SNR(A)$ can be used regardless of the frequency characteristics of D while SNR_{uni32} cannot used for the frequency characteristics with a steep slope such as the frequency characteristic of (f).

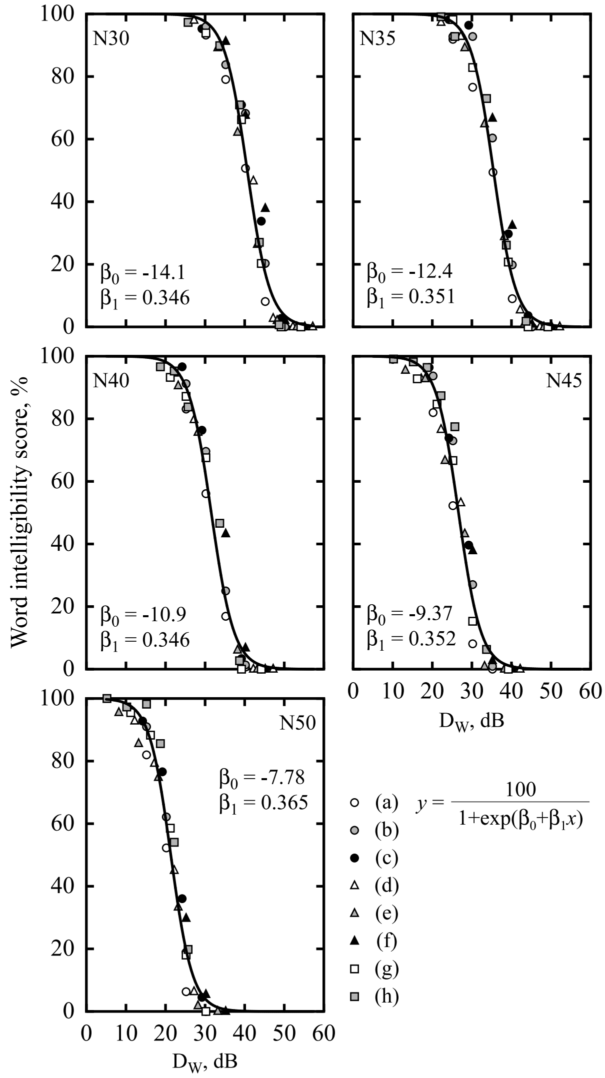

Figure 7: Word intelligibility scores as a function of D_W for each presentation level of the additional noise. NX indicates the result for the additional noise level of X dBA. Different symbols represent different frequency characteristics of D. The solid curves represent logistic regression curves for each additional noise level. β_0 and β_1 are the intercept and the regression coefficient of D_W, respectively.

Evaluation from sound insulation performance and background noise level

Being able to estimate word intelligibility scores from a number of variables that can be changed is useful for assessing or designing speech privacy or security in a confidential room. Regression analyses were performed to obtain an estimate equation of word intelligibility scores which includes sound insulation performance and background noise level as independent variables.

The weighted level difference (D_W)[6] was obtained for each D to use it as a variable that corresponds to sound insulation performance. D_W was obtained using the 1/1 octave method, and the reference curve was moved in 0.1 dB steps to increase resolution of sound insulation performance. Figure 7 represents the word intelligibility scores as a function of D_W for each presentation level of the additional noise. Different symbols represent different frequency characteristics of D. Solid curves represent logistic regression curves for each additional noise level.

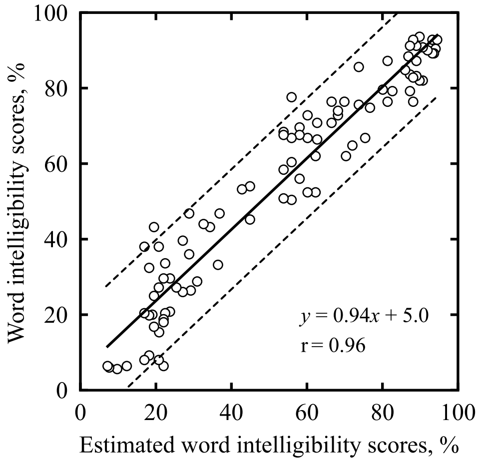

Figure 8: The relationship between word intelligibility scores and estimated scores based on a multiple logistic regression analysis (see Eq. 2). The solid line represents a linear regression line between the two scores, and the dashed lines represent 95% prediction intervals.

The word intelligibility scores were fitted in logistic regression curves for each additional noise level, regardless of the frequency characteristic of D and the additional noise level. It should be noted that the regression coefficients of D_W (β_1) for each additional noise level were almost equal to each other. The estimated standard errors of β_1 were around 0.01 for all additional noise level, and t-tests based on the estimated standard errors demonstrated that there were no statistically significant differences between all pairs of β_1 ($p<0.05$). This means that the effects of D_W and the additional noise level on the scores were independent.

A multiple logistic regression analysis was performed to obtain a estimate equation of the scores from D_W and the additional noise level, which independently affected the scores. Equation 2 is a regression equation obtained from the analysis, where WI: Word intelligibility scores (%), D_W: Weighted level difference (dB), and L_N: Additional noise level (dBA).

$$WI = 100/(1+\exp(-24.3+0.352D_W+0.335L_N)) \quad (2)$$

Figure 8 represents the relationship between the scores and those estimated from Eq. 2. The relationship between the scores and those estimated from Eq. 2 was very similar to that for $SNR(A)$ (see. Fig. 5(B)). The correlation coefficient (r=0.96) was slightly higher than that for $SNR(A)$ (r=0.94), and 95% prediction interval was smaller by around $\pm3\%$ than that for $SNR(A)$ at the average of the estimated scores. Therefore, it is concluded that word intelligibility scores can be estimated from D_W and background noise level using Eq. 2 with the same or higher accuracy relative to that of $SNR(A)$, regardless of the frequency characteristics of D.

Equal-intelligibility contours

Figure 9 shows the equal-intelligibility contours based on Eq. 2. This chart enables us to easily estimate required D_W and A-weighted background noise level to achieve a desired word intelligibility score in the next room. For example, when D_W is 31 dB and A-weighted background noise level is 40 dB, the word intelligibility score in an adjacent room will be 50%. If the score is desired to be reduced to 10%, D_W should be increased to 37 dB, or the background noise level should be increased to 47 dBA.

It should be noted that Fig. 9 applies to only L_{S1}, that is, the

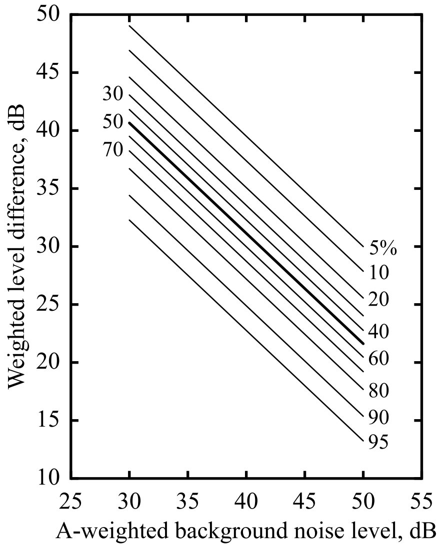

Figure 9: Equal-intelligibility contours as a function of the weighted level difference and A-weighted background noise level.

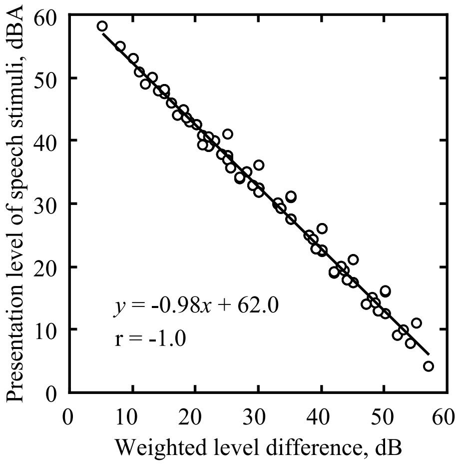

Figure 10: The relationship between the presentation level of speech stimuli and the weighted level difference.

speech level at the position of person B in Fig. 1, of 58 dBA. If different L_{S1} have to be assumed, the required D_W and A-weighted background noise level should be corrected. Figure 10 represents the presentation level of speech stimuli that is corresponding to the speech level at the position of person C (L_{S2}) in Fig. 1 as a function of D_W. The correlation between L_{S2} and D_W was very high (r$\approx$-1), and the slope of the regression line was about -1. This means that a 1 dB increase of L_{S2} can be replaced as a 1 dB decrease of D_W. Needless to say, L_{S1} is linked to L_{S2} and a 1 dB increase of L_{S1} causes a 1 dB increase of L_{S2}. Therefore, the required D_W should be increased by the same amount of the increase of L_{S1} from 58 dBA.

Figure 11 is a modified version of Fig. 9. The ordinate axis is replaced as “Weighted level difference - ΔL, dB” , and ΔL is defined as L_{S1} - 58. This chart would be able to apply to any L_{S1}. The fact that the increase of vocal effort causes change of the frequency characteristic of speech would not affect the estimate accuracy of Fig. 11, because the relationship between the scores and D_W is not affected by the frequency characteristics of D as shown in Fig. 7 and the change of the frequency characteristic of speech can be replaced as that of D.

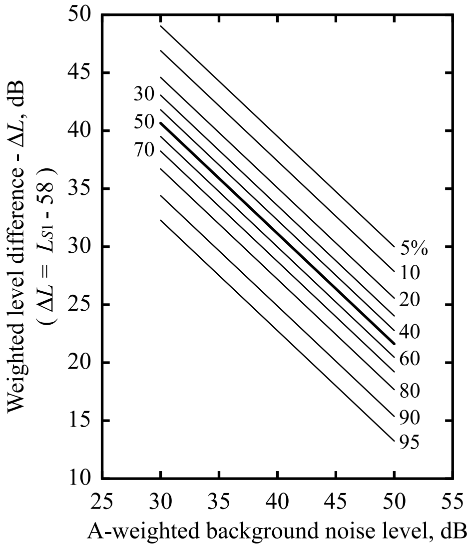

Figure 11: Equal-intelligibility contours which is modified to apply any L_{S1}.

SUMMARY

In the present study, word intelligibility tests were performed to clarify the relationship among sound insulation performance, background noise level, and the degree of conversation leakages. The results of tests and analyses are summarized as follows.

(1) A-weighted speech-to-noise ratio and SNR_{uni32} can estimate word intelligibility scores with high accuracy, regardless of the frequency characteristics and the absolute values of the sound pressure level difference, and background noise level. However, SNR_{uni32} underestimates the scores when the frequency characteristic of the sound pressure level difference has a steep slope.

(2) The weighted level difference and background noise level independently affect word intelligibility scores. A multiple logistic regression analysis with the scores as a dependent variable, and the weighted level difference and background noise level as independent variables shows that the scores can be estimated from the two independent variables with the same or higher accuracy relative to $SNR(A)$.

(3) Equal-intelligibility contours, that can easily show the weighted level difference and background noise level required to achieve a certain level of word intelligibility scores, were obtained from the result of the multiple logistic regression analysis. Furthermore, the modified contours to apply any speech level or vocal effort in rooms where confidential conversation takes place were also suggested.

ACKNOWLEDGMENT

This research project was partially supported by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (B), 20360262, and Ono Acoustics Research Fund.

REFERENCES

1 W. J. Cavanaugh, W. R. Farrell, P. W. Hirtle, and B. G. Watters. Speech privacy in buildings. *J. Acoust. Soc. Am.*, 34:475–492, 1962.
2 ANSI S3.5-1969. *American National Standard Methods for the Calculation of the Articulation Index.* 1969.
3 B. N. Gover and J. S. Bradley. Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms. *J. Acoust. Soc. Am.*, 116:3480–3490, 2004.
4 H. K. Park, J. S. Bradley, and B. N. Gover. Evaluating airborne sound insulation in terms of speech intelligibility. *J. Acoust. Soc. Am.*, 123:1458–1471, 2008.
5 ASTM E413. *Classification for Rating Sound Insulation.*
6 ISO 717-1:1996. *Acoustics - Rating of sound insulation in buildings and of building elements - Part 1: Airborne sound insulation.* 1996.
7 J. S. Bradley, M. Apfel, and B. N. Gover. Some spatial and temporal effects on the speech privacy of meeting rooms. *J. Acoust. Soc. Am.*, 125:3038–3051, 2009.
8 W. O. Olsen. Average speech levels and spectra in various speaking/listening conditions: A summary of the Peason, Bennett, & Fidell (1977) report. *Am. J. Audiol.*, 7:21–25, 1998.
9 H. Tachibana, Y. Hamada, and F. Sato. Loudness evaluation of sounds transmitted through walls: Basic experiment with artificial sounds. *J. Sound & Vib.*, 127:499–506, 1988.
10 Z. Maekawa and P. Lord. *Environmental and Architectural Acoustics.* E&FN SPON, London, UK, 1993.
11 JIS A 1419-1:2000. *Acoustics – Rating of sound insulation in buildings and of building elements – Part 1: Airborne sound insulation.* 2000 (in Japanese).
12 S. Sakamoto, Y. Suzuki, S. Amano, K. Ozawa, T. Kondo, and T. Sone. New lists for word intelligibility test based on word familiarity and phonetic balance. *J. Acoust. Soc. Jpn.*, 54:842–849, 1998 (in Japanese).
13 H. Kuttruff. *Room Acoustics (3rd ed.).* Elsevier, London, UK, 1991.