Listening difficulty as a subjective measure for evaluation of speech transmission performance in public spaces

Morimoto, Masayuki
Sato, Hiroshi
Kobayashi, Masaaki

(Citation)
Journal of the Acoustical Society of America,116(3):1607-1613

(Issue Date)
2004-09

(Resource Type)
journal article

(Version)
Version of Record

(URL)
https://hdl.handle.net/20.500.14094/90000852

Listening difficulty as a subjective measure for evaluation of speech transmission performance in public spaces

Masayuki Morimoto[a)]
Environmental Acoustics Laboratory, Faculty of Engineering, Kobe University, Rokko, Nada, Kobe 657-8501 Japan

Hiroshi Sato
Graduate School of Engineering, Tohoku University, Aoba, Sendai 980-8579 Japan

Masaaki Kobayashi
Environmental Acoustics Laboratory, Faculty of Engineering, Kobe University, Nada, Kobe 657-8501 Japan

(Received 7 January 2004; revised 24 May 2004; accepted 31 May 2004)

Listening difficulty ratings, using words with high word familiarity, are proposed as a new subjective measure for the evaluation of speech transmission in public spaces to provide realistic and objective results. Two listening tests were performed to examine their validity, compared with intelligibility scores. The tests included a reverberant signal and noise as detrimental sounds. The subject was asked to repeat each word and simultaneously to rate the listening difficulty into one of four categories: (1) not difficult, (2) a little difficult, (3) fairly difficult, and (4) extremely difficult. After the tests, the four categories were reclassified into, not difficult [response (1)] and some level of difficulty, (the other 3 responses). Listening difficulty is defined as the percentage of the total number of responses indicating some level of difficulty [i.e. not (1)]. The results of two listening tests demonstrated that listening difficulty ratings can evaluate speech transmission performance more accurately and sensitively than intelligibility scores for sound fields with higher speech transmission performance. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1775276]

PACS numbers: 43.55.Hy, 43.71.Gv [MK] [NX] Pages: 1607–1613

I. INTRODUCTION

The acoustical design of rooms often involves optimizing speech transmission performance since speech communication is a basic function of most public spaces such as classrooms, meeting rooms, auditoriums, railway stations, airports, etc. In the past, many studies have tried to evaluate the performance of these spaces using both physical and subjective measures.

Some physical measures are based on the useful-to-detrimental ratio, which was developed by Lochner and Burger[1] and used in a simplified form as U50.[2] Clarity[3] and Deutlicheit[4] are closely related to useful-detrimental ratios but without the influence of ambient noise. Bradley[5] showed that measured values of these indices and the more recent STI (Speech Transmission Index)[6] are closely related and can be estimated from each other. Meanwhile, several studies have developed subjective measures, such as the Rhyme test,[7] the SPIN (Speech Perception in Noise) test,[8] the SRT[9] (Speech Recognition Threshold) and so on. Among them, SRT is defined as the presentation level of speech necessary for listeners to recognize the speech 50% correctly. Because we are interested in the near perfect scores required for good speech transmission and minimal communication difficulty, SRT values are not relevant to this study.

Conventional subjective measures can be roughly classified into two groups. One is the syllable articulation test using mono-syllables, or tri-syllables, etc. A defect of syllable articulation tests is that it is difficult to imagine speech transmission performance because test syllables used in listening tests lack the reality of real speech. The other choice is intelligibility tests using words and sentences which assess both phonetic and linguistic components of the process of speech perception. Listeners indicate the percentage of words or sentences that they recognize. In these tests it is easy to imagine the actual speech transmission performance. However, the intelligibility score depends on the choice of words and sentences used in the listening tests.[8] Thus, each of the syllable articulation tests and intelligibility tests using words and sentences has both merits and disadvantages. Furthermore, even if the intelligibility is perfect, the speech transmission performance is not always satisfactory. Although a 0 dB useful-to-detrimental sound ratio corresponds to above 90% intelligibility,[5] this is not always adequate for ideal speech communication in spaces. In response to these problems, subjective ratings of the "easiness" of speech recognition were presented by Sato *et al.*[10] as an alternative approach. However, "easiness" ratings result in a relative scale and one cannot say how easy is good enough or when a low rating corresponds to unacceptably bad acoustical conditions. In addition, "easiness" lacks objectivity and would be influenced by personal preferences. It should probably be considered as related to speech quality.

From an engineering point of view, it is important to obtain a physical measure, which can evaluate the speech transmission performance. To achieve this purpose, it is necessary to compare the subjective measures with objective measures of realistic situations that could be a reference for

[a)]Electronic mail: mrmt@kobe-u.ac.jp

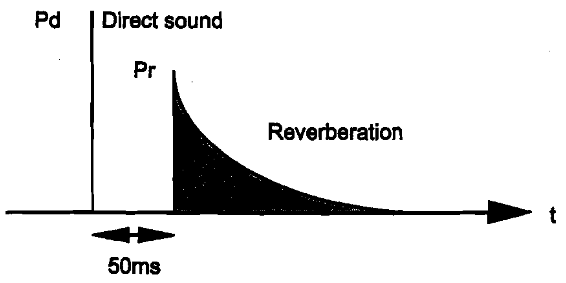

FIG. 1. Impulse response of the test sound field. Pd and Pr is a sound pressure level of a direct sound and a reverberation signal, respectively.

judging whether or not the physical measure is useful for evaluating speech transmission performance.

To solve the problems mentioned above, the present authors propose "listening difficulty" using words with high "word familiarity," as a subjective measure, which is adequate to imagine the speech transmission performance and does not depend on the test words. "Word familiarity" is a subjective variable which indicates how familiar native speakers are with the word. The familiarity of about eighty thousand Japanese words has been rated from 1.0 (minimum) to 7.0 (maximum) by Amano *et al.*[11] Another reason why high familiarity words are used is to minimize the effect of the cognitive process on the difficulty rating since the main object of interest in this study is only the speech transmission performance of spaces and including these with public address systems. Moreover, the reason why "difficulty" is introduced, instead of "easiness" is that people more frequently experience listening difficulty in daily life. In other words, "difficulty" has reality and objectivity and matches our normal feelings more than "easiness."

Our purpose in this study is to examine the validity of "listening difficulty," and to compare it with intelligibility scores.

II. LISTENING TEST I: LISTENING DIFFICULTY AND INTELLIGIBILITY FOR SOUND FIELDS WITH REVERBERATION

First, the validity of listening difficulty is examined by using a simple sound field that consisted of a direct sound and a reverberant signal as detrimental sound.

A. Listening test

1. Test sound fields and test words

Figure 1 shows schematically the impulse response of the test sound field. The sound field was composed of a direct sound followed by a reverberant decay. The reverberant decay started at 50 ms after the direct sound. As Bradley *et al.*[12] reported, increased early reflection energy arriving at the listener within the first 50 ms after the direct sound has the same effect on speech intelligibility scores as an equal increase in the direct sound energy. In this study, simulated sound fields are used including only the direct sound component as useful sound energy and reverberant signals as detrimental sound energy. The parameters were reverberation time and the sound pressure ratio of the initial amplitude of the reverberant decay to the direct sound amplitude described as Pr/Pd. The reverberation time was set at 0.5, 0.9, 2.0 and 6.0 s. The Pr/Pd was set at 0.5, 0.25, and 0.1. Table I shows the eight kinds of sound field used in the listening test.

TABLE I. Eight kinds of sound field used in the first listening test.

	Sound field							
Parameter	A	B	C	D	E	F	G	H
Reverberation time (s)	0.5	0.9	0.9	0.9	2.0	2.0	2.0	6.0
Pr/Pd	0.5	0.5	0.25	0.1	0.5	0.25	0.1	0.5

Two word lists with two different levels of word familiarity in Japanese[11,13] were used in this test. The word familiarity, categorized from 1.0 to 7.0, as mentioned above was divided into four levels. One was the word list with the level 4 familiarity corresponding to the categories from 7.0 to 5.5, which was used for the subjective measure proposed here. The other was the word list with the level 3 familiarity corresponding to the categories from 5.5 to 4.0. This word list was added to the test in order to determine the effect of word familiarity on the difficulty ratings. In this paper, the former and the latter are called "high familiarity" and "low familiarity," respectively. Each word list included 50 phonetically balanced words. Each word has four syllables.

2. Subjects

A total of 122 subjects were used in the test. They varied in age from 19 to 31 years old. They did not report any known hearing impairment. Thus the subjects were not carefully selected listeners, but were more representative of general young listeners with normal hearing.

3. Procedure

In the test, five words of the 50 words with each familiarity were presented to each subject for each of eight different sound fields. Thus, each subject listened to a set of 80 test signals in total ($5\ \text{words} \times 2\ \text{familiarities} \times 8\ \text{sound fields}$) in a random order. Moreover, each word was presented to a subject only once. A different set of 80 test signals was presented to 122 subjects so that all 800 test signals ($50\ \text{words} \times 2\ \text{familiarities} \times 8\ \text{sound fields}$) might be presented in the whole of the listening test. At the beginning of the test, 12 test signals, every other of six words with each familiarity for each of three test sound fields A, E, and H, respectively, were added as a practice trial. The interval for presentation of each test signal was 9 s. The total length of the test was about 19 min for each subject.

The test signals were presented to two subjects at a time from a loudspeaker in an anechoic room. The listening positions were located at 3 m distance from the loudspeaker and at ± 30 degrees from the central axis of the loudspeaker. The frequency response of the loudspeaker was flat to within ± 5 dB from 100 to 10 kHz at both listening positions. The sound pressure levels of the test signals were set at 65.0 dB A, fast, peak on the average of all test words at listening positions when only the direct sound was presented. The range of the levels of the test words was within ± 2.0 dB A.

Each subject was asked to repeat by writing down each word as they listened and simultaneously to rate the listening

TABLE II. Four categories for listening difficulty.

(1) Not difficult
(2) A little difficult
(3) Fairy difficult
(4) Extremely difficult

difficulty into one of the four categories shown in Table II. The object of measuring listening difficulty is classifying conditions as either to "no-difficulty" or "difficulty." However, it was difficult for listeners to identify the boundary between "no-difficulty" and "difficulty" and to answer in terms of one of the two categories. The solution employed in the listening test was to ask subjects to answer in terms of one of four categories described in Table II and then to calculate the difficulty rating with the boundary between "no-difficulty" and "difficulty" set in analysis which is going to be explained later.

B. Results and discussion

1. Definition of listening difficulty

Figure 2 indicates the results of the listening test for sound fields with reverberation. Panel (a) is the result for the high familiarity words and the panel (b) is that for the low familiarity words. The abscissa indicates the sound field and the ordinate indicates the score as a percentage. The number of responses was 610 (5 words per sound field $\times$122 subjects). Open circles are the intelligibility scores. Open triangles are the listening difficulty ratings "1." Closed circles are the sum of the difficulty ratings "1" and "2." Crosses are the sum of the difficulty ratings "1," "2," and "3." Both of these results indicate that open circles do not coincide with open triangles, but coincide well with closed circles. This means that even if the subjects can repeat words correctly, they feel them difficult to listen to in some cases. In other words, even if the intelligibility is 100%, the speech transmission performance in these cases cannot always be regarded as satisfactory. The speech transmission performance could be regarded as perfect when the percentage of responses for listening difficulty is zero.

Accordingly, listening difficulty is defined as a subjective measure of speech transmission performance that indicates the percentage of the sum of the difficulty responses "2," "3," and "4" (i.e., except "1"—not difficult). From this definition and the results mentioned above, it can be considered that the listening difficulty can evaluate the speech transmission performance more accurately than intelligibility scores.

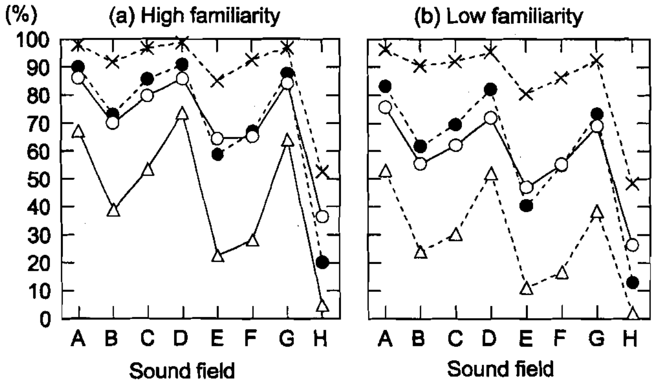

FIG. 2. Results of the first listening test. Left panel: high familiarity. Right panel: low familiarity. Open circles indicate the intelligibility scores. Open triangles, closed circles and crosses indicate the responses of difficulty (1), (1)+(2), and (1)+(2)+(3), respectively.

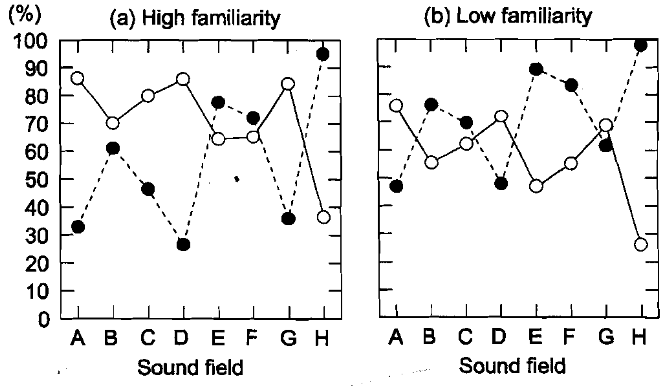

FIG. 3. Listening difficulty ratings and the intelligibility scores. Left panel: high familiarity. Right panel: low familiarity. Open circles indicate listening difficulty ratings and closed circles indicate the intelligibility scores.

2. Relation between listening difficulty and intelligibility

Figure 3 shows listening difficulty ratings and the intelligibility scores for each test sound field. Closed circles indicate the listening difficulty ratings and open circles indicate the intelligibility scores. For both word familiarities, it can be seen that there is a negative correlation between listening difficulty and intelligibility. Figure 4 shows the correlation between listening difficulty and intelligibility. Each point represents a different sound field. As indicated in Fig. 4, correlation coefficients indicate a strong negative correlation between them for both familiarities. The regression equations were obtained such that "x" is the listening difficulty and "y" is the intelligibility. For the high familiarity, the coefficient of x is 0.653 that is half of the coefficient of y. This means that the listening difficulty can evaluate the speech transmission performance more sensitively than the intelligibility. Also the same tendency can be seen for the low familiarity words, although the coefficient is a little larger than for the high familiarity words. In summary, it is

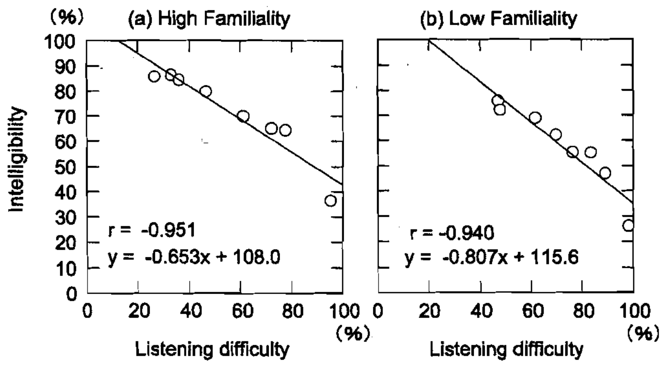

FIG. 4. The relation between listening difficulty and intelligibility. Left panel: high familiarity. Right panel: low familiarity. In the resulting regression equations, "x" is the listening difficulty and "y" is the intelligibility.

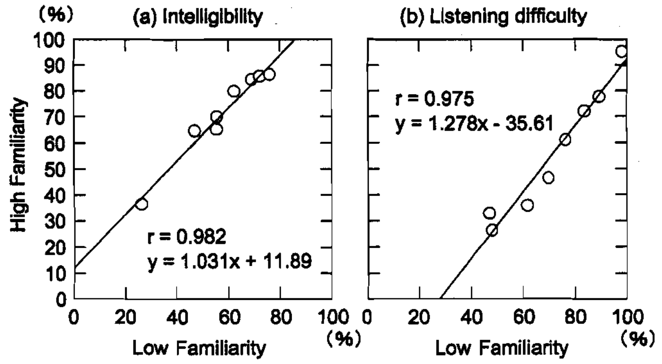

FIG. 5. Relation between the high familiarity and the low familiarity. Left panel: intelligibility. Right panel: listening difficulty. In the resulting regression equations, "x" is the low familiarity and "y" is the high familiarity.

concluded that listening difficulty ratings can evaluate speech transmission performance more accurately and sensitively than intelligibility scores in this test region.

3. Effects of word familiarity on listening difficulty and intelligibility

Figure 5 shows the correlation between results for words with the high and the low familiarities for intelligibility [panel (a)] and listening difficulty [panel (b)]. Each point represents a different sound field. It can be seen that when the word familiarity becomes lower, intelligibility becomes lower and the listening difficulty becomes higher. The correlation coefficients and the regression equations between results for the high familiarity and the low familiarity were obtained. As shown in Fig. 5, there is a strong correlation between both the intelligibility and listening difficulty for low and high familiarity words. Furthermore, the coefficients of x in the regression equations are almost equal. This means that the effects of word familiarity on intelligibility and listening difficulty do not depend on the test sound fields.

III. LISTENING TEST II: LISTENING DIFFICULTY AND INTELLIGIBILITY FOR SOUND FIELDS WITH REVERBERATION AND NOISE

In this second experiment, the validity of listening difficulty is examined by using sound fields consisting of a direct sound, plus a reverberation signal and noise as detrimental sound. Bradley[5] demonstrated that an optimum speech transmission performance can be obtained when the signal-to-noise (SN) ratio is at least 15 dB A for reverberation times less than 0.5 s. In this section, the effects of noise level and reverberation time on the listening difficulty are investigated and compared with the effects on intelligibility scores to determine acceptable SN ratios for good speech transmission.

A. Listening test

1. Test sound fields and test words

The impulse response of the test sound field is the same as shown in Fig. 1. The Pr/Pd is kept constant at 0.1. The sound pressure levels of the test signals were kept constant at 55.0 dB A, slow, peak for the average of all test words at listening positions when only the direct sound was presented. The range of the levels of test words was within ± 2.0 dB A. The parameters were reverberation time and background noise level. The reverberation time was set at 0.5 and 2.0 s. The background noise level was set at 0 (no noise was presented from the loudspeaker), 10, 25, 40, 55, and 60 dB A. The ambient noise level in the experimental room was 17.8 dB A. The level corresponds to the speech signal (direct sound) to noise ratio (SN ratio) of infinity (no additional noise), 45, 30, 15, 0, and -5 dB A, respectively. Hence the total number of sound fields used in the listening test was twelve (2 reverberation times $\times$ 6 background noise levels). The power spectrum of noise used as a background noise is the same as that found by Hoth[14] for each 1/1 octave band from 125 Hz to 8 kHz.

Six word lists (300 words) with the high familiarity in Japanese were used in the test. One of the lists was also used in the first test.

2. Subjects

Thirteen university students with normal hearing acted as the subjects in the test. They did not serve as subjects for the first test.

3. Procedure

In the test, 25 words of 300 words were presented to each subject for each of 12 different sound fields. Thus each subject listened to a set of 300 test signals in total (25 words $\times$ 12 sound fields). Moreover, each word was presented to a subject only once. Three hundred test signals were arranged in a random order and divided into six units of 50 words. Six units were presented separately. Different sets consisting of 300 test signals were presented to 13 subjects so that all 3600 test signals (300 words $\times$ 12 sound fields) might be presented during the whole of the listening test. The duration of each word was around 700 ms. The duration of each test signal depended on reverberation time. The duration of a background noise for each test signal was around 7 s depending on the duration of the test signal. The background noise was presented in advance of the test signal by 135 ms and was ended after each test signal ended. The interval for presentation of background noise was 6 s. During the interval no noise was presented. The total length of each unit was about 10 min.

The test signals were presented to two subjects at a time from a loudspeaker in an anechoic room as in the first test. The listening positions and the loudspeaker used for the test were the same as those in the first test.

Each subject was asked to repeat by writing down each test word as they listened and to rate the listening difficulty into one of the four categories shown in Table II.

B. Results and discussion

Figure 6 indicates the results of the listening tests for sound fields with reverberation and noise. The abscissa indicates the noise level and SN ratio and the ordinate indicates the score as a percentage. The number of responses was 325 (25 words per sound field $\times$ 13 subjects). Circles indicate the

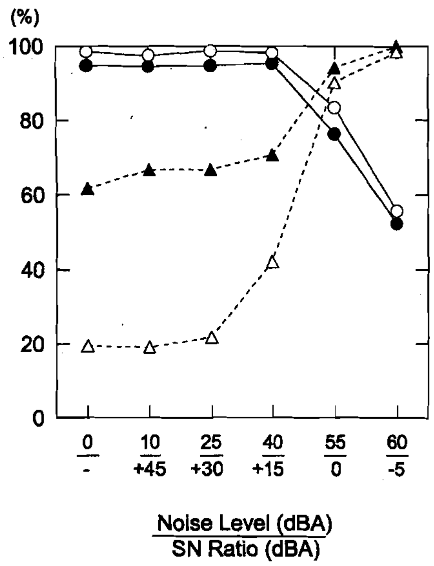

FIG. 6. Listening difficulty ratings and the intelligibility scores as a function of noise level and speech signal to noise ratio. Open and filled symbols indicate scores for the reverberation time of 0.5 and 2.0 s, respectively. Circles represent the intelligibility and triangles indicate the listening difficulty. A 0 dB A noise level means no additional noise was presented from the loudspeaker. The ambient noise level in the experimental room was 17.8 dB A.

intelligibility and triangles indicate the listening difficulty. Open and filled symbols represent scores for the reverberation time of 0.5 and 2.0 s, respectively.

First, let us discuss both types of scores for the sound fields with noise of less than 40 dB A (SN ratio > 15 dB A). The intelligibility is close to 100%, regardless of the reverberation time. On the other hand, the listening difficulty for the reverberation time of 0.5 s is better than that for 2.0 s. The difference is about 20% for noise of 40 dB A (SN ratio = 15 dB A) and 40% for less than 25 dB A (SN ratio > 30 dB A). Furthermore, the listening difficulty approaches a fixed value for noise of less than 25 dB A (SN ratio > 30 dB A). The value is about 20% for the reverberation time of 0.5 s and about 60% for 2.0 s. The listening difficulty never reaches 0% with reverberation present, but the intelligibility increases to 100% as the noise level decreases.

Next, let us discuss both scores for the sound fields with noise of more than 55 dB A (SN ratio < 0 dB A). Listening difficulty is affected by neither the noise level nor the reverberation time. Listening difficulty reaches 100% for noise of more than 55 dB A (SN ratio < 0 dB A). On the other hand, intelligibility is clearly affected by the noise level, but not by the reverberation time. The intelligibility is about 80% and 50% for noise levels of 55 dB A (SN ratio = 0 dB A) and 60 dB A (SN ratio = −5 dB A), respectively.

These results suggest that listening difficulty is not always high when background noise is present. Listening difficulty ratings can evaluate speech transmission performance more sensitively and accurately than intelligibility scores when the background noise is lower than 40 dB A (SN ratio > 15 dB A). Conversely, intelligibility scores can evaluate speech transmission performance more sensitively than listening difficulty ratings when background noise is higher than 40 dB A (SN ratio < 15 dB A). Namely, listening difficulty ratings can evaluate speech transmission performance more sensitively than the intelligibility when the performance is high, and *vice versa*.

Finally let us discuss a necessary SN ratio for good speech transmission when the speech level is 55 dB A. Bradley[2] reported that the desired ratio is at least 15 dB A for a reverberation time of 0.5 s. The results of this listening test also demonstrated that the intelligibility is almost 100% under the same condition as that mentioned above. However, there is a difference of 20% in the listening difficulty between SN ratios of 15 and 30 dB A for the same reverberation time. In other words, listening difficulty evaluates the speech performance more strictly than does intelligibility. As a result, for the minimum listening difficulty, the necessary SN ratio is higher than 15 dB A, when the speech level is 55 dB A.

Additionally, there is probably an inverse correlation between useful-to-detrimental energy and listening difficulty but further study in various S/N and acoustical conditions is required to find a clear relation between them.

IV. FURTHER DISCUSSION ON SUBJECTIVE MEASURES OF SPEECH TRANSMISSION

The relation between subjective measures of speech transmission and the performance of speech transmission of sound fields is discussed based on the results obtained in the present paper.

Figure 7 illustrates different relations between them. Panel (a) presents the ideal relation between them. The subjective value monotonically increases as the performance is advanced. The subjective value reaches the lowest when the performance is minimum and it reaches the highest when the performance is maximum.

At present, however, no subjective measure demonstrates such a relationship as shown in panel (a).

At present, the most common listening test material for the subjective evaluation of speech transmission is an intelligibility test such as the rhyme test. Most previous[2,13,15] papers have demonstrated relations as illustrated in panel (b). Namely, that intelligibility scores can differentiate between the performance of sound fields where the performance is relatively low. On the other hand, it cannot clearly identify differences in the performance of sound fields where the performance is relatively high, because the intelligibility reaches a plateau near to 100%.

Meanwhile, in the present paper we propose listening difficulty ratings as a subjective measure for an evaluation of the speech transmission performance and demonstrate that (1) Listening difficulty and intelligibility scores show a highly negative correlation and listening difficulty ratings can evaluate speech transmission performance more sensitively than intelligibility scores when the performance is high, and intelligibility scores can evaluate speech transmission performance more sensitively than difficulty ratings when the performance is low. (2) Intelligibility is high even when the listening difficulty is close to 100%. (3) Listening difficulty changes significantly and does not reach 0% even if the intelligibility is 100%. These results yield the relation

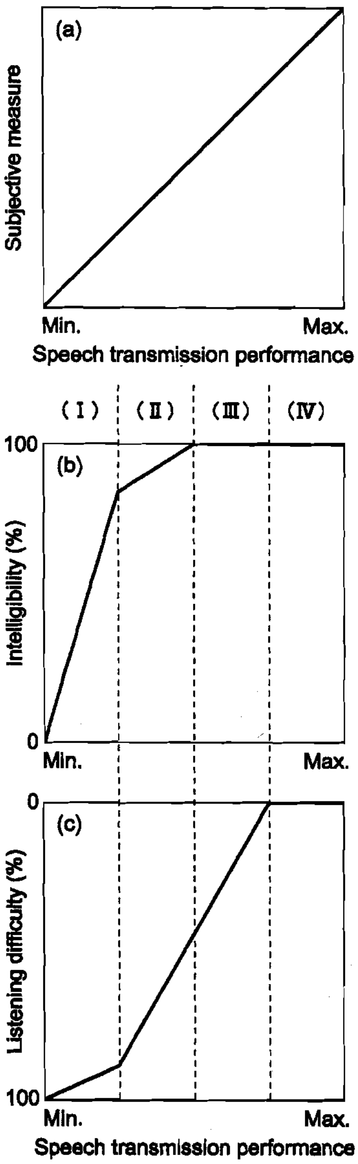

FIG. 7. Schematic relations between subjective measures of speech transmission and the performance of speech transmission. Panel (a) presents the ideal relation. Panel (b) illustrates the relation between the performance and the intelligibility. Panel (c) shows the relation between the performance and the listening difficulty.

between listening difficulty and speech transmission performance as shown in panel (c).

In summary, the relation between intelligibility and listening difficulty can be explained reasonably by dividing the speech transmission performance into four ranges from (I) to (IV) as shown in Fig. 7. In the range (I) where the speech transmission performance is the worst, the performance can be evaluated sensitively by using the intelligibility. On the other hand, the listening difficulty is around 100% in this range and it cannot describe clearly the differences in performance. This condition corresponds to the test sound fields with noise levels of 55 and 60 dB A used in the second test. In the range (II), the performance is a little higher, and listening difficulty and intelligibility show a highly negative correlation. Furthermore, listening difficulty can evaluate the speech transmission performance more accurately and sensitively than does intelligibility. The eight test sound fields used in the first test are in this range. In the range (III), where the performance is higher, the intelligibility reaches 100% and cannot differentiate among the conditions. On the other hand, listening difficulty can evaluate the performance sensitively. This condition corresponds to the test sound fields with noise levels of less than 40 dBA used in the second test.

Actually, it is usually not required to evaluate the performance of sound fields belonging to the range (I), since actual sound fields to be evaluated, such as classrooms, meeting rooms, railway stations, auditoriums, airports etc., have higher speech transmission performance than this range (I). Therefore, it can be concluded that listening difficulty is more suitable to evaluate speech transmission performance in public spaces than intelligibility scores, because difficulty can evaluate the performance more sensitively and accurately than intelligibility scores in the ranges (II) and (III).

Sound fields in the range (IV), in which nobody feels listening difficulty (0% difficulty) and of course in which the intelligibility is 100%, must be evaluated by using other measures such as the speech quality which may have higher sensitivity than listening difficulty ratings in this range.

V. CONCLUSIONS

In the present paper we have proposed listening difficulty ratings using words with the maximum familiarity as a subjective measure that can evaluate speech transmission performance realistically and objectively. Two kinds of listening tests were carried out to validate listening difficulty ratings compared with intelligibility scores. Sound fields with reverberation and others with reverberation and noise were used as test sound fields in the first and the second listening tests, respectively. The results of the two listening tests clarified that (1) Listening difficulty ratings can evaluate speech transmission performance more accurately and sensitively than intelligibility scores for sound fields with a higher speech transmission performance. (2) Conversely, intelligibility scores can evaluate the performance more sensitively than listening difficulty ratings for the sound fields with lower performance. (3) It is possible to estimate the listening difficulty for words of one word familiarity from those of another word familiarity in these test regions. (4) A necessary speech signal to noise ratio for good speech transmission is at least 15 dB A in terms of intelligibility, while it is at higher than 15 dB A in terms of listening difficulty. In conclusion, listening difficulty ratings are more suitable to evaluate actual sound fields, such as classrooms, meeting rooms, railway stations, auditoriums, airports etc., than intelligibility scores.

ACKNOWLEDGMENTS

The authors acknowledge NTT Basic Research Laboratories and Research Institute of Electrical Communication of Tohoku University for providing the phonetically balanced word lists based on word familiarity. The authors are grateful to Dr. John S. Bradley for his fruitful discussions. A part of this study was carried out as a project of the Sub Working Group for Developing Evaluation and Design Standards of Speech Transmission Quality at the Architectural Institute of

Japan. The research project was partially supported by the research grant from Kajima Foundation and by the Twenty-first Century Center of Excellence (COE) Program "Design Strategy towards Safety and Symbiosis of Urban Space" awarded to the Graduate School of Science and Technology, Kobe University. The Ministry of Education, Culture, Sports, Science and Technology of Japan sponsored the Program.

[1] J. P. A. Lochner and J. F. Burger, "The intelligibility of speech under reverberant conditions," Acustica **11**, 195–200 (1961).

[2] J. S. Bradley, "Predictors of speech intelligibility in rooms," J. Acoust. Soc. Am. **80**, 837–845 (1986).

[3] W. Reichardt and U. Lehmann, "Optimierung raumeindruck und durchsichtigkeit von musikdarbietungen durch auswertung von impulsschalltests," Acustica **48**, 174–185 (1981).

[4] R. Thiele, "Richtungsverteilung und zeitfolge der schallruckwurfe in raumen," Acustica **3**, 291–302 (1953).

[5] J. S. Bradley, "On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility," J. Acoust. Soc. Am. **106**, 1820–1828 (1999).

[6] T. Houtgast and H. J. M. Steeneken, "The modulation transfer function in room acoustics as a predictor of speech intelligibility," Acustica **28**, 66–73 (1973).

[7] G. Fairbanks, "Test of phonetic differentiation: The Rhyme Test," J. Acoust. Soc. Am. **30**, 596–600 (1958).

[8] D. N. Kalikow, K. N. Stevens, and L. L. Elliott, "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," J. Acoust. Soc. Am. **61**, 1337–1351 (1977).

[9] R. Plomp and M. Mimpen, "Speech-reception threshold for sentences as a function of age and noise level," J. Acoust. Soc. Am. **66**, 1333–1342 (1979).

[10] H. Sato, H. Yoshino, and M. Nagatomo, "Relationship between speech transmission index and easiness of speech perception in reverberatory fields," J. Acoust. Soc. Am. **103**, 2999 (1998).

[11] http://www.ais.riec.tohoku.ac.jp/lab/wordlist/index-j.html (in Japanese).

[12] J. S. Bradley, H. Sato, and M. Picard, "The importance of early reflections for speech intelligibility in rooms," J. Acoust. Soc. Am. **113**, 3233–3244 (2003).

[13] 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).

[14] D. F. Hoth, "Room noise spectra at sbscribers's telephone locations," J. Acoust. Soc. Am. **12**, 499 (1941).

[15] J. S. Bradley, "Speech intelligibility studies in classrooms," J. Acoust. Soc. Am. **80**, 846–854 (1986).