<table>
<thead>
<tr>
<th>タイトル (Title)</th>
<th>Relationship between listening difficulty and acoustical objective measures in reverberant sound fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>著者 (Author(s))</td>
<td>Sato, Hayato / Morimoto, Masayuki / Sato, Hiroshi / Wada, Megumi</td>
</tr>
<tr>
<td>掲載誌・巻号・ページ (Citation)</td>
<td>Journal of the Acoustical Society of America, 123(4):2087-2093</td>
</tr>
<tr>
<td>刊行日 (Issue date)</td>
<td>2008-04</td>
</tr>
<tr>
<td>資源タイプ (Resource Type)</td>
<td>Journal Article / 学術雑誌論文</td>
</tr>
<tr>
<td>版区分 (Resource Version)</td>
<td>publisher</td>
</tr>
<tr>
<td>権利 (Rights)</td>
<td>©2008 Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America. The following article appeared in Journal of the Acoustical Society of America 123(4), 2087-2093 and may be found at <a href="http://dx.doi.org/10.1121/1.2885750">http://dx.doi.org/10.1121/1.2885750</a></td>
</tr>
<tr>
<td>DOI</td>
<td>10.1121/1.2885750</td>
</tr>
<tr>
<td>JaLCDOI</td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://www.lib.kobe-u.ac.jp/handle_kernel/90004033">http://www.lib.kobe-u.ac.jp/handle_kernel/90004033</a></td>
</tr>
</tbody>
</table>
Relationship between listening difficulty and acoustical objective measures in reverberant sound fields

Hayato Sato\textsuperscript{a} and Masayuki Morimoto

Environmental Acoustics Laboratory, Graduate School of Engineering, Kobe University, Rokko, Nada, Kobe 657-8501, Japan

Hiroshi Sato

Institute for Human Science & Biomedical Engineering, National Institute of Advanced Industrial Science and Technology, Tsukuba Central, Tsukuba, Ibaraki 305-8568, Japan

Megumi Wada

Environmental Acoustics Laboratory, Graduate School of Engineering, Kobe University, Rokko, Nada, Kobe 657-8501, Japan

(Received 6 August 2007; revised 29 January 2008; accepted 3 February 2008)

The previous work [Morimoto et al., J. Acoust. Soc. Am. 116, 1607–1613] showed that listening difficulty ratings can be used to evaluate speech transmission performance more exactly and sensitively than intelligibility. Meanwhile, speech transmission performance is usually evaluated using acoustical objective measures, which are directly associated with physical parameters of room acoustic design. However, the relationship between listening difficulty ratings and acoustical objective measures was not minutely investigated. In the present study, a total of 96 impulse responses were used to investigate the relationship between listening difficulty ratings and several objective measures in unidirectional sound fields. The result of the listening test showed that (1) the correlation between listening difficulty ratings and speech transmission index (STI) is the strongest of all tested objective measures, and (2) A-weighted $D_{50}$, $C_{50}$, and center time, which are obtained from the impulse responses passed through an A-weighted filter, also strongly correlate with listening difficulty ratings, and their correlations with listening difficulty ratings are not statistically different from the correlation between listening difficulty ratings and STI.

© 2008 Acoustical Society of America. [DOI: 10.1121/1.2885750]

PACS number(s): 43.55.Hy, 43.71.Gv [NX] Pages: 2087–2093

I. INTRODUCTION

Evaluation of speech transmission performance is essential for designing rooms for speech communication. Ideally, speech transmission performance should be subjectively evaluated by listeners, because the performance indicates how accurately and comfortably speech information is transmitted to listeners. Morimoto et al.\textsuperscript{1} suggested listening difficulty ratings as a subjective measure to evaluate the speech transmission performance of rooms. Listening difficulty ratings are the percentage of responses that indicate some level of difficulty in listening to the most familiar words. Sato et al.\textsuperscript{2} reported that listening difficulty is a more appropriate evaluation tool than word intelligibility tests for conditions with speech-to-noise ratios between 0 and 15 dBA that are commonly found in everyday life. Therefore, it is advisable that the design and evaluation of rooms for speech communication be carried out on the basis of listening difficulty ratings.

However, objective measures are usually used to evaluate speech transmission performance, because they are directly associated with the physical parameters of room acoustic design, and it takes much time to obtain subjective measures. Therefore, the relationship between subjective and objective measures needs to be clarified in advance to convert objective measure to subjective measure, and vice versa.

Background noise and reverberant sounds are critical factors in determining speech transmission performance in rooms. Useful-to-detrimental ratio proposed by Lochner and Burger,\textsuperscript{3} and speech transmission index (STI) proposed by Houtgast and Steeneken\textsuperscript{4} can estimate the combined effects of the two critical factors on speech intelligibility. The two objective measures have different concepts in estimating the effect of reverberation sounds on speech intelligibility. The concept of useful-to-detrimental ratio is that the energy of the direct sound and reflected sounds can be divided into useful energy and detrimental energy for speech intelligibility according to the interval from the direct sound. Meanwhile, the concept of STI is based on the reduction of the modulation transfer function (MTF) due to reverberant sounds. Although the concept of STI is different from that of useful-to-detrimental ratio, Bradley\textsuperscript{5} reported that the correlation coefficients between speech intelligibility and the two objective measures are essentially equivalent.

The relationship between listening difficulty ratings and objective measures was investigated in the previous studies.\textsuperscript{1,2,6} Morimoto et al.\textsuperscript{1} and Kobayashi et al.\textsuperscript{6} reported that listening difficulty ratings increased with increasing reverberation time and decreasing speech-to-noise ratio. Sato

\textsuperscript{a}Author to whom correspondence should be addressed. Electronic mail: hayato@kobe-u.ac.jp.
et al.\textsuperscript{2} reported that a logistic curve could be used to explain the relationship between listening difficulty ratings and $U_{50(A)}$, that is, A-weighted useful-to-detrimental ratio with 50 ms early time interval. However, reverberant sounds used in the previous studies\textsuperscript{1,2,6} were artificial and simple ones, while reverberant sounds in real rooms have complex properties and vary widely. In addition, the number of reverberant sound fields used in the previous studies was not enough to investigate the relationship between listening difficulty ratings and objective measures.

In the present study, the relationship between listening difficulty ratings and objective measures is investigated focusing on the effect of time sequences of sound reflections. A total of 96 reverberant sound fields that cover a wide variety of acoustic conditions were tested. These sound fields were unidirectional, and were simulated by convolving test signals with monaural impulse responses measured in real rooms.

II. METHOD

A. Test words and sound fields

A hundred Japanese words were used as test words. Test words were selected from the word lists proposed by Sakamoto et al.\textsuperscript{7} to be most familiar to young adults. Each test word has four syllables. Test words were spoken by a female speaker, and recorded in an anechoic room.

A total of 96 impulse responses were used as test sound fields. Seventy-one of them were selected from the impulse response database compiled by the Speech Communication Research Working Group of the Architectural Institute of Japan.\textsuperscript{8} The database included 966 digitized impulse responses measured at various types of rooms in Japan, for example, a meeting room, an auditorium, a concert hall, a film theater, and a sports area. The impulse responses, which had been measured using a one-point omni-directional microphone, were used in the present study. The impulse responses were selected to have a wide variety of acoustic conditions.

The other 25 impulse responses were artificially created to cover less and much reverberant sound fields that are lacking in number in the database. Artificial impulse responses were created by reverberating or trimming monaural impulse responses in the database\textsuperscript{8} using software (SEKD Sampler) on a personal computer. Figure 1 represents the numbers of real and artificial impulse responses used in the listening test as a function of STI. Open and closed bars respectively represent the numbers of real and artificial impulse responses. The STIs of artificial impulse responses were controlled in order to flatten the distribution of the number of impulse responses as a function of STI.

B. Participant

Fifty-five young adults (male: 27, female: 28) participated in the listening test. They were university students in their twenties. The results of pure-tone audiometry show that all participants had normal hearing sensitivity. All participants had the experience of participating in listening tests on evaluating speech transmission performance.

C. Procedure

Each of the test words was preliminarily convolved with each of the test impulse responses on a personal computer. Thus, a total of 9600 test signals (100 words $\times$ 96 impulse responses) were prepared. The test signals were divided into 100 different sets consisting of 96 test signals.

Moreover, four reverberation-free words were added to each set. Sato et al.\textsuperscript{9} reported that the reproducibility of listening difficulty ratings improved when a reverberation-free and quiet sound field and an extremely reverberant and noisy sound field were included in the test sound fields as reference signals. The four reverberation-free words were expected to work as the highest-quality reference signals, and the lowest-quality reference signals, that is, extremely reverberant sound fields, were already included in the test sound fields in the present study. Therefore, each set consisted of 96 test signals and four highest-quality reference signals. Each set contained all test words and impulse responses, and combinations of the test words and the test impulse responses were different in different sets.

Each participant listened to four different sets of test signals. A total of 220 sets of test signals (55 participants $\times$ 4 sets) were presented throughout the listening test. Each set was used twice or thrice. The listening test was divided into eight sessions each of which allows participants to listen to 50 test signals. Test signals were presented in random order, except that the first signal in each session was always the highest-quality reference signal. The interval for the presentation of each signal was 10 s. It took about 7 min. to complete each session.

The test signals were unidirectionally presented from a loudspeaker (Fujitsu Ten, TD512) at a distance of 1.5 m in front of the participant in an anechoic room. The frequency characteristics of the loudspeaker were flat within $\pm$5 dB in the range from 100 Hz to 10 kHz. The sound pressure level of each test signal was measured at the position of the center of the subject’s head using a sound level meter, which was set at A-weighted and slow response. The differences among the peak levels of the test signals were within $\pm$1.5 dBA, and the level averaged over all test signals was set at 65 dBA. No additional noise was added to the test signals.

![FIG. 1. Numbers of real and artificial impulse responses (IR) used in the listening test as a function of STI in steps of 0.05. Open and closed bars respectively represent frequencies of real and artificial impulse responses.](image-url)
TABLE I. Categories of listening difficulty (see Ref. 1).

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Not</td>
<td>difficult</td>
</tr>
<tr>
<td>2</td>
<td>A little</td>
<td>difficult</td>
</tr>
<tr>
<td>3</td>
<td>Fairly</td>
<td>difficult</td>
</tr>
<tr>
<td>4</td>
<td>Extremely</td>
<td>difficult</td>
</tr>
</tbody>
</table>

Each participant was asked to take dictation of each test signal as they listened using katakana characters (Japanese phonograms) and to simultaneously rate the listening difficulty into one of the four categories shown in Table I. Before the listening test, each participant listened to ten signals as an exercise. The signals for exercises were made by convolving ten words that were not included in the 100 test words with ten different impulse responses used in the listening test so that the signals for exercises contained the widest range of the test sound fields.

D. Acoustical objective measures

A total of six types of acoustical objective measures were tested in the present study. The tested objective measures were speech transmission index (STI),\textsuperscript{4} Deutlichkeit (D\textsubscript{50}),\textsuperscript{10} an early-to-late ratio with the fixed boundary at 50 ms from the direct sound (C\textsubscript{50}),\textsuperscript{11} center time (t\textsubscript{c}),\textsuperscript{12} reverberation time (RT), and early decay time (EDT). STI was obtained from the impulse responses according to the original method documented in a review by Houtgast and Steeneken.\textsuperscript{13} The objective measures except STI were obtained from the impulse responses passed though 1 octave band filters with center frequencies from 125 Hz to 8 kHz. In addition, a two-octave-band filter that covers octave bands of 500 Hz and 1 kHz, and an A-weighted filter were used for a trial, because these filters are often used to evaluate acoustical environments. Furthermore, the frequency weighting sum of D\textsubscript{50} and that of C\textsubscript{50} were also obtained using the weighting factors for STI\textsuperscript{13} and speech intelligibility index\textsuperscript{14} (1/1 octave band method, female) as a trial. The slope of bandpass filters was −48 dB/oct. RT and EDT were, respectively, obtained from the slopes of the reverberation energy decay curve from −5 to −35 dB and from −5 to −15 dB using Shoreder’s method.\textsuperscript{15}

III. RESULTS AND DISCUSSION

A. Relationship between word intelligibility scores and listening difficulty ratings

Word intelligibility scores and listening difficulty ratings for each impulse response were obtained from the collective results of the listening test for all participants. This means that the number of samples for each impulse response was 220. The word intelligibility score is the percentage of the test signals written down correctly. The listening difficulty rating is the percentage of the sum of listening difficulty ratings evaluated from “2” to “4” in Table I. Note that listening difficulty ratings decrease when speech transmission performance improves, in contrast to word intelligibility scores.

The listening difficulty rating for the reverberant-free reference signal was 1.3%, and the maximum listening difficulty rating was 98%. These results indicated that the sound fields used in the present study included both high-end and low-end reference signals, and that the listening difficulty ratings obtained in the present study will have high reproducibility.

Figure 2 represents the relationship between word intelligibility scores and listening difficulty ratings. Word intelligibility scores decreased with increasing listening difficulty ratings. However, word intelligibility scores varied over a range of only 27% (i.e., from 73 to 100%), while listening difficulty ratings varied over almost the full range (from 2 to 98%). This behavior corresponds to those observed in the previous studies.\textsuperscript{1,2,6}

B. Relationship between listening difficulty ratings and objective measures

Generally speaking, the relationship between a subjective value and an objective value is nonlinear, and a cumulative normal distribution function or a logistic curve is often used to approximate a subjective-objective function. As reported by Sato et al.,\textsuperscript{2} this relationship seems to be similar to the case of the relationship between listening difficulty ratings and objective measures. In the present study, a cumulative normal distribution function was used to approximate listening difficulty ratings as a function of each objective measure. Listening difficulty ratings were converted into standard scores (z scores)\textsuperscript{16} to simplify the approximation using a linear regression analysis. The z scores of −1, 0, and 1, respectively, correspond to the listening difficulty ratings of 16, 50, and 84%.

Linear regression analyses between z scores of listening difficulty ratings and each objective measure were performed. The listening difficulty ratings that were less than 5% (z < −1.64) and more than 95% (z > 1.64) were omitted to prevent their excessive effect on linear regression analyses. A total of 16 test sound fields were omitted in the following analyses.


1. Correlation coefficient

Table II represents correlation coefficients between $z$ scores of listening difficulty ratings and each of the objective measures. The correlation between $z$ scores of listening difficulty ratings and STI was the strongest of all objective measures ($r=-0.97$). The differences between the absolute value of correlation coefficient for STI and the other absolute values of correlation coefficients were statistically tested using Fisher’s $z$ transformation. The lower limit of the absolute value of correlation coefficient that is not statistically different from that for STI is calculated to be 0.937 ($p<0.05$). This means that the correlation coefficients between $-0.97$ and 0.937 are statistically weaker than that for STI. Asterisks on the right-hand side of the correlation coefficients in Table II denote that the correlation is statistically weaker than that for STI. The correlations for RT and EDT were statistically weaker than that for STI for all cases. It was common to observe for $D_{50}, C_{50}$, and $t_s$ that the correlations for the A-weighted filter were not statistically weaker than that for STI. The correlations for 1–4 kHz were instead stronger than those for 250 Hz, 500 Hz, and 8 kHz regardless of the objective measures. These results indicate that listening difficulty is more strongly affected by reverberant sounds within octave bands from 1 to 4 kHz than other octave bands. The A-weighted function relatively increases sound energy within octave bands from 1 to 4 kHz, and decreases other frequency sounds that are not very important for listening difficulty ratings. Therefore, A-weighted objective measures seem to strongly correlate with listening difficulty ratings. The mid-two-octave band (500 Hz–1 kHz) objective measures, the frequency weighting sum of $D_{50}$, and that of $C_{50}$ also showed strong correlations, which were almost the same as those for A-weighted objective measures.

2. Scatter diagram, linear regression line and prediction interval

Figure 3 represents scatter diagrams for STI and A-weighted objective measures. A linear regression line and 95% prediction intervals are also shown in each panel. Linear regression equations for other frequencies are shown in the Appendix. The prediction intervals for EDT(A) and RT(A) were around ±0.9σ at the $z$ score of listening difficulty ratings of zero. Meanwhile, the prediction intervals for other objective measures were around half as small as those for EDT(A) and RT(A). The prediction intervals at the $z$ score of listening difficulty ratings of zero are ±0.43σ for STI, ±0.57σ for $D_{50}(A)$, ±0.56σ for $C_{50}(A)$ and ±0.51σ for $t_s(A)$. The differences between the prediction interval for STI and that for other objective measures ranged from 0.08 to 0.14σ. This also indicates that the estimation accuracies of $D_{50}(A), C_{50}(A)$ and $t_s(A)$ are not inferior to that of STI. The relationship between $z$ scores of listening difficulty ratings and $t_s(A)$ showed a logarithmic function rather than a linear function. However, the correlation coefficient between $z$ scores of listening difficulty ratings and the base-10 logarithm of $t_s(A)$ ($r=-0.96$) was not significantly increased from that for $t_s(A)$.

Note that the test sound fields used in the present study were unidirectional sound fields. It is unclear whether the results of the present study are applicable in real sound fields or not at this time. Further studies will clarify this problem. Considering a masking level difference due to spatial aspects of reverberant sounds, listening difficulty ratings in real sound fields might improve relative to those estimated from monaural objective measures. Even if that were true, the estimation error would be at least on the safe side.

C. Evaluation of speech transmission performance using listening difficulty and objective measures

Figure 2 clearly demonstrates that listening difficulty ratings can present the difference between speech transmission performances that word intelligibility scores cannot. Moreover, Fig. 3 shows that $z$ scores of listening difficulty ratings were linearly proportional to STI in the range of STI from 0.3 to 0.9. This demonstrates that listening difficulty

---

**TABLE II.** Correlation coefficients between the $z$ scores of listening difficulty ratings and each of the objective measures. The objective measures except STI were calculated using the impulse responses passed through 1/1 octave bandpass filters for each center frequency, a two-octave bandpass filter that covers octave bands of 500 Hz and 1 kHz, and an A-weighted filter. The frequency weighting sum of $D_{50}$ and that of $C_{50}$ were obtained using the weighting factors for STI (see Ref. 13) and speech intelligibility index (Ref. 14). (SII).

<table>
<thead>
<tr>
<th>Frequency weighting</th>
<th>STI</th>
<th>$D_{50}$</th>
<th>$C_{50}$</th>
<th>$t_s$</th>
<th>EDT</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 Hz</td>
<td>$-0.74^*$</td>
<td>$-0.75^*$</td>
<td>0.81*</td>
<td>0.76*</td>
<td>0.67*</td>
<td></td>
</tr>
<tr>
<td>250 Hz</td>
<td>$-0.86^*$</td>
<td>$-0.85^*$</td>
<td>0.92*</td>
<td>0.77*</td>
<td>0.72*</td>
<td></td>
</tr>
<tr>
<td>500 Hz</td>
<td>$-0.88^*$</td>
<td>$-0.88^*$</td>
<td>0.92*</td>
<td>0.80*</td>
<td>0.81*</td>
<td></td>
</tr>
<tr>
<td>1 kHz</td>
<td>$-0.94^*$</td>
<td>$-0.93^*$</td>
<td>0.94</td>
<td>0.81*</td>
<td>0.83*</td>
<td></td>
</tr>
<tr>
<td>2 kHz</td>
<td>$-0.93^*$</td>
<td>$-0.92^*$</td>
<td>0.95</td>
<td>0.82*</td>
<td>0.83*</td>
<td></td>
</tr>
<tr>
<td>4 kHz</td>
<td>$-0.92^*$</td>
<td>$-0.93^*$</td>
<td>0.96</td>
<td>0.80*</td>
<td>0.85*</td>
<td></td>
</tr>
<tr>
<td>8 kHz</td>
<td>$-0.88^*$</td>
<td>$-0.88^*$</td>
<td>0.92*</td>
<td>0.72*</td>
<td>0.82*</td>
<td></td>
</tr>
<tr>
<td>Two-oct. bandpass</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500-1 kHz</td>
<td>$-0.93^*$</td>
<td>$-0.93^*$</td>
<td>0.94</td>
<td>0.81*</td>
<td>0.82*</td>
<td></td>
</tr>
<tr>
<td>A-weighted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency weighting</td>
<td>STI</td>
<td>$-0.97$</td>
<td>$-0.94$</td>
<td>$-0.94$</td>
<td>$-0.94$</td>
<td>$-0.94$</td>
</tr>
<tr>
<td>SII</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistically different from STI ($p<0.05$)
ratings can present the difference between most speech transmission performances of room acoustics that appear in real situations.

Objective measures are used to evaluate the speech transmission performance of rooms in various stages of acoustic design. Before construction of rooms, estimation of RT from volume and sound absorbing power is usually performed on the assumption of a diffuse sound field. However, the correlation between listening difficulty and reverberation time is significantly weaker than in the case of other objective measures. Therefore, it is better to consider RT as a simple but rough predictor of listening difficulty. Meanwhile, impulse responses are needed to obtain objective measures other than RT. Bistafa and Bradley\(^\text{18}\) suggested the formulation of \(D_{50}\) and \(C_{50}\) in a diffuse sound field by introducing an ideal impulse response consisting of the direct sound and exponential decay that start at the same time. Houtgast and Steeneken\(^\text{13}\) also suggested the formulation of MTF in a diffuse sound field based on the same ideal impulse response. Barron’s revised theory\(^\text{19}\) also enables us to obtain \(D_{50}\) and \(C_{50}\) from RT and source-receiver distance. These would be more accurate predictors of listening difficulty than RT at sound receiving points where the direct sound is dominant.

Impulse responses can be obtained from measurements in scale model or computer simulations before constructing rooms. These methods are useful for optimizing speech transmission performance by arranging reflective or absorptive materials on room surfaces. In this manipulation, information on whether each reflected sound is useful or detrimental for speech perception is directly related to the arrangement of room surfaces. The concept of an early-to-late ratio more clearly provides this information than that of STI. Considering that A-weighted \(D_{50}\) and \(C_{50}\) strongly correlate with listening difficulty ratings and that the correlations are not statistically different from the correlation for STI, A-weighted \(D_{50}\) and \(C_{50}\) are more useful than STI from the viewpoint of the acoustical design of rooms.

After construction of rooms, impulse responses can be measured in the constructed rooms. STI can estimate listening difficulty most accurately from impulse responses of rooms. Therefore, it is better to use STI to conclusively determine listening difficulty in the constructed rooms. A-weighted \(D_{50}\), \(C_{50}\), and \(t_e\) are simple alternatives for such use. Figure 4 represents the relationship between listening difficulty ratings and other measures to evaluate speech transmission performance on the basis of the linear regres-

---

**FIG. 3.** Z score of listening difficulty rating as a function of each of the objective measures. A solid line and dashed lines in each panel respectively represent a linear regression line and 95% prediction intervals.
sponds to an STI of 0.5, and the boundary between intelligi-

able, a listening difficulty rating of 50% roughly corre-

sponds to an STI of 0.5, and the boundary between intelligi-

ability ratings of “Fair” and “Good” corresponds to a listening
difficulty rating of around 30%.

IV. CONCLUSION

The relationship between listening difficulty ratings and
acoustical objective measures was investigated using the
monaural impulse responses measured at various types of
rooms in Japan, focusing on the effect of time sequences of
sound reflections. The results of listening tests and analyses
clarified that (1) the correlation between listening difficulty
ratings and STI is the strongest of all tested objective mea-
sures, and (2) A-weighted D_{50}, C_{50}, and t_s, which are cal-
culated from the impulse responses passed through an
A-weighted filter, also strongly correlated with listening
difficulty ratings, and their correlations with listening difficulty
ratings are not statistically different from the correlation be-
tween listening difficulty ratings and STI.

Note that the results of the present study were obtained
in quiet sound fields. The effect of background noise needs to
be investigated to determine the best objective measure for
estimating listening difficulty ratings in noisy sound fields,
such as public spaces.

ACKNOWLEDGMENTS

This research project was partially supported by the re-
search and development grant of the Japan Institute of
Construction and Engineering, 03011, 2004, and by the Japan
Society for the Promotion of Science, Grant-in-Aid for Sci-
entific Research (B), 16360292, 2004–2006.

APPENDIX: LINEAR REGRESSION EQUATIONS
BETWEEN Z SCORES OF LISTENING DIFFICULTY
RATINGS AND ACOUSTICAL OBJECTIVE
MEASURES

TABLE A1. Linear regression equations and correlation coefficients
between z scores of listening difficulty ratings (y) and objective measures (x)
for octave bands of 500 Hz, 1 kHz, and 2 kHz, and for two-octave bands of
500 Hz and 1 kHz.

<table>
<thead>
<tr>
<th>Objective measure (x)</th>
<th>Linear regression equation</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_{50}(500)</td>
<td>y = -2.93x + 1.08</td>
<td>-0.88</td>
</tr>
<tr>
<td>D_{50}(1k)</td>
<td>y = -3.01x + 1.27</td>
<td>-0.94</td>
</tr>
<tr>
<td>D_{50}(2k)</td>
<td>y = -2.84x + 1.14</td>
<td>-0.93</td>
</tr>
<tr>
<td>D_{50}(500-1k)</td>
<td>y = -3.03x + 1.25</td>
<td>-0.93</td>
</tr>
<tr>
<td>C_{50}(500)</td>
<td>y = -0.134x - 0.38</td>
<td>-0.88</td>
</tr>
<tr>
<td>C_{50}(1k)</td>
<td>y = -0.137x - 0.24</td>
<td>-0.93</td>
</tr>
<tr>
<td>C_{50}(2k)</td>
<td>y = -0.126x - 1.27</td>
<td>-0.92</td>
</tr>
<tr>
<td>C_{50}(500-1k)</td>
<td>y = -0.140x - 0.27</td>
<td>-0.93</td>
</tr>
<tr>
<td>t_s(500)</td>
<td>y = 0.49x - 1.39</td>
<td>0.92</td>
</tr>
<tr>
<td>t_s(1k)</td>
<td>y = 0.56x - 1.32</td>
<td>0.94</td>
</tr>
<tr>
<td>t_s(2k)</td>
<td>y = 0.36x - 1.34</td>
<td>0.95</td>
</tr>
<tr>
<td>t_s(500-k)</td>
<td>y = 0.77x - 1.36</td>
<td>0.94</td>
</tr>
<tr>
<td>EDT(500)</td>
<td>y = 0.793x - 1.56</td>
<td>0.80</td>
</tr>
<tr>
<td>EDT(1k)</td>
<td>y = 0.773x - 1.59</td>
<td>0.81</td>
</tr>
<tr>
<td>EDT(2k)</td>
<td>y = 0.804x - 1.62</td>
<td>0.82</td>
</tr>
<tr>
<td>EDT(500-1k)</td>
<td>y = 0.780x - 1.58</td>
<td>0.81</td>
</tr>
<tr>
<td>RT(500)</td>
<td>y = 0.902x - 1.62</td>
<td>0.81</td>
</tr>
<tr>
<td>RT(1k)</td>
<td>y = 0.889x - 1.65</td>
<td>0.83</td>
</tr>
<tr>
<td>RT(2k)</td>
<td>y = 0.957x - 1.73</td>
<td>0.83</td>
</tr>
<tr>
<td>RT(500-1k)</td>
<td>y = 0.900x - 1.65</td>
<td>0.82</td>
</tr>
</tbody>
</table>

M. Morimoto, Hi. Sato, and M. Kobayashi, “Listening difficulty as a sub-
jective measure for evaluation of speech transmission performance in

Hi. Sato, J. Bradley, and M. Morimoto, “Using listening difficulty ratings

J. P. A. Lochner and J. F. Burg, “The influence of reflections on audito-

I. Houtgast and H. J. M. Steeneken, “The modulation transfer function in
room acoustics as a predictor of speech intelligibility,” Acustica 28, 66–73
(1973).

J. S. Bradley, “Relationships among measures of speech intelligibility in

level to minimize listening difficulty in public spaces,” J. Acoust. Soc.
Am. 121, 251–256 (2007).

S. Sakamoto, Y. Suzuki, S. Amano, K. Ozawa, T. Kondo, and T. Sone,
“New lists for word intelligibility test based on word familiarity and pho-

transmission performance with impulse response in rooms,” J. Environ.

Ha. Sato, M. Morimoto, A. Hakamada, M. Kobayashi, and Hi. Sato,
“Learning effect and reduction of context effect on listening difficulty for
words,” Memoirs of Graduate School of Science and Technology, Kobe

R. Theile, “Richtungsverteilung und Zeitfolge der Schallrückwürfe in
Räumen,” Acustica 3, 291–302 (1953). (Title in English: “Directional
distribution and time sequence of sound reflections in rooms.”)


R. Küfer, “Zur Gewinnung von Einzahlkriterien bei Impulsmessungen in
der Raumakustik,” Acustica 21, 370–372 (1969). (Title in English: “Ex-
traction of single-number criteria from impulse measurements of room
acoustics.”)
14 ANSI S3.5-1997, Methods for the Calculation of the Speech Intelligibility Index (American National Standards Institute, New York, 1997).