Relationship between speech transmission index and easiness of speech perception in reverberation

Hiroshi Sato*, Hiroshi Yoshino* and Muneshige Nagatomo**

* Dept. of Architecture, Faculty of Eng., Tohoku University, Sendai 980-8579 and
** Kajima Technical Research Institute, Choufu 182-0036, Japan

Abstract: The concept of Easiness of Speech Perception (ESP) is a kind of auditory impression, and different from accuracy of speech recognition. This study consists of two experiments that conducted in reverberatory fields. The first experiment was carried out in electric reverberation fields. The other was done in fields simulated actual auditoriums. It was found that ESP correlated with STI except for cases with reverberation and/or binaural factor effects.

INTRODUCTION

The indicators of speech intelligibility such as the syllable articulation test and the word intelligibility test, express accuracy of speech recognition in a sound field. By using the concept of Easiness of Speech Perception (ESP), the speech transmission quality of a space can be evaluated directly.

Cases are presented that ESP show a different tendency from syllable articulation (1). ESP is useful scale to distinguish the difference of the speech transmission quality of rooms in which the score of speech intelligibility tests are same but auditory impressions are different.

The Speech Transmission Index (STI), proposed by T. Houtgast & H.J.M. Steeneken (2), was used as an indicator of sound fields and is compared with ESP in this study.

EXPERIMENTS

The first of the two experiments had five cases. Reverberation Time (T) and one of the other five parameters as shown in FIGURE 1 varied as indicated in TABLE 1 for each cases. The parameters, which were fixed in the case, were set at the standard condition shown in TABLE 1. Reverberation were added by the digital reverberator (YAMAHAREV7).

The other experiment was done in the sound fields that used the impulse responses measured in auditoriums. The single corn loud speaker placed at center of stage radiated the time stretched pulse for measuring the impulse response. A monaural microphone and a dummy head were used for monaural listening and binaural listening, respectively. Twenty impulse responses were used in the experiment. Half of them were measured in the auditoriums with Sound Reinforcement (SR) system. The peak level of speech without reverberation was set for 70dBA at the listening point.

Psychological magnitude of ESP was obtained by the Sheffe's paired comparison method. The stimulus used for the measurement of ESP had a duration time of 5 seconds. An interval between two stimuli was 8 seconds for direct sound. All pairs of stimuli were arranged in random order and separated by an interval of 10 seconds for direct sound.

Five subjects (22-28 years old) were used in the first experiment, and three of them were also used in the second. Two trials were carried out per person at each condition.

TABLE 1. Conditions of Parameters varied in the first experiment

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverberation Time (s)</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>Peak Level of Direct Sound (dBA)</td>
<td>60, 63*, 66, 69</td>
</tr>
<tr>
<td>Pd to Pr Ratio (Pd / Pr *)</td>
<td>2, 4, 8*, 16</td>
</tr>
<tr>
<td>Initial Delay Time (ms)</td>
<td>30, 50*, 80, 110</td>
</tr>
<tr>
<td>Frequency Characteristics of Reverberation Time</td>
<td>A, B, C, D**</td>
</tr>
<tr>
<td>Frequency Characteristics of Reverberation Sound</td>
<td>A, B, C, D***</td>
</tr>
</tbody>
</table>

* A: Low-Up (+7.5% increase at 1kHz), B: High-Down (-26% decrease at 8kHz), C: High-Down, D: Low-Up, E: Flat

** A: Low Boost (+5dB at 1kHz) & High Cut (-5dB at 8kHz), B: High Cut, C: Low Boost, D: Flat

* * * * : Standard Condition

FIGURE 1. Parameters varied in the first experiment (T=2.0s, 2 oct. bands at 500 to 1kHz)
RESULTS AND DISCUSSION

FIGURE 2 shows the result of the first experiment. It is indicated that ESP has a good correlation with STI. The correlation coefficient of them are 0.99 except for the case where Pd/Pr was changed. Therefore, STI can predict ESP in simple reverberatory fields in case where Pd/Pr is 4 or more.

FIGURE 3 shows the result of the second experiment. There are two correlations, one for using the SR system, the one for not using it. The both correlation are good (R>0.9) in monaural simulation fields, but correlation is worse for the case with the SR system in binaural simulation. The reason for the phenomenon described above is that frequency response of the SR system and the location and number of speakers are not considered by STI and may exert an influence on ESP.

FIGURE 2. Relationship between psychological scale of ESP and STI in reverberatory fields of the first experiment

FIGURE 3. Relationship between psychological scale of ESP and STI in simulated fields of the second experiment

ACKNOWLEDGMENT

This research was supported in part by a grant from the Sound Technology Promotion Foundation in Japan.

REFERENCES