Free-Field Virtual Psychoacoustics and Hearing Impairment: 
Paper ICA2016-123

Study on evaluation of speech intelligibility focusing on speech privacy

Hyojin Lee (a), Souhei Tsujimura (b), Shinichi Sakamoto (c)

(a) Institute of Industrial Science, the University of Tokyo, Japan, leehj@iis.u-tokyo.ac.jp
(b) Railway Technical Research Institute, Japan, tsouhei@rtri.or.jp
(c) Institute of Industrial Science, the University of Tokyo, Japan, sakamo@iis.u-tokyo.ac.jp

Abstract

Recently, concerns about speech privacy are increasing continuously in Japan. In order to achieve the speech privacy, proper evaluation method of the speech privacy is required. Speech intelligibility has a high relationship with the speech privacy. However, the relationship between impression of oral information leakage and the speech intelligibility is not clear in Japan. In this study, the speech intelligibility test and categorical rating tests for investigating the impression of oral information leakage were measured according to the change in the speech levels. In addition, influences of background noise on the speech intelligibility test and the categorical rating tests were also examined.

Keywords: speech privacy, evaluation method, speech intelligibility test, categorical rating test
Study on evaluation of speech intelligibility focusing on speech privacy

1 Introduction

Speech privacy [1] to avoid oral information leakage in healthcare facilities has become an important issue in Japan. In order to realize requirements of the speech privacy, standardization of the speech privacy also has been discussed [2]. On evaluation of oral information leakage, speech intelligibility tests are useful. Among them, one of the most general methods in Japan is a word intelligibility test by using words with the same word-familiarity. The word intelligibility test is a simple and convenient evaluation method to obtain quantitative values. However, relationships between the speech intelligibility test scores and degrees of audibility and understanding of speech are still not clear. In this study, we conducted two speech intelligibility tests and two categorized subjective judgement tests on degrees of audibility and understanding of speech by varying its speech levels. In addition, influence of background noise on evaluation of oral information leakage was investigated. The tests were carried out in an anechoic room with two loud speakers.

2 Experimental method

The evaluation of the oral information leakage was carried out by using the two intelligibility test categorized subjective judgement tests. A general description of the tests is as follows.

1) Word intelligibility test (WI test): Subjects are exposed to one sound condition with 10 words, and requested to dictate the word given during the test.

2) Sentence intelligibility test (SI test): Short sentences explaining specific words were presented as the test sound, and the subjects were requested to dictate the specific words. Ten sentences were selected for each sound condition. The speech intelligibility test reflects a human ability to infer a whole context of story out of limited number of words. In this test, subjects were allowed to guess context.

3) Categorized subjective judgement test: After the subjects dictated the words of the SI test, the subjects are requested to reply to the following two questions. The first was the impression of "degree of audibility of speech". The second was the impression of "degree of understanding of speech". The subjects answered based on five categories; 1.Not at all, 2.Slightly, 3.Moderately, 4.Very, 5.Extremely.

2.1 Sound reproduction system

The tests were conducted in a 7 cubic meters anechoic room. The test sounds were reproduced through two loudspeakers (TANNOY, T12) 2 m away from a subject (Figure 1). The Speech was always played over a loudspeaker positioned at the front of the subject. The background noise was played over a loudspeaker positioned either at the front, above the subject depending on the test condition.
2.2 Test condition

Table 1 shows level setting of the speech and the background noise. The levels of the speech were changed in 5 dB steps. As the background noise, pink noise was reproduced. The presented levels were 45 dB. The test conditions were 16 in total and the presented order were randomized.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Speech Loudspeaker position</th>
<th>Speech Level [L_{Aeq}, dB]</th>
<th>Background noise Loudspeaker position</th>
<th>Background noise Level [L_{Aeq}, dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Front</td>
<td>20, 25, 30, 35, 40, 45</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Case 2</td>
<td>Front</td>
<td>25, 30, 35, 40, 45</td>
<td>Above</td>
<td>45</td>
</tr>
<tr>
<td>Case 3</td>
<td>Front</td>
<td>25, 30, 35, 40, 45</td>
<td>Front</td>
<td>45</td>
</tr>
</tbody>
</table>
As the speech materials, a clear female voice made by software was used. For the WI test, four mora words with the same word-familiarity were chosen from the “NTT Database Series: Lexical Properties of Japanese” [3]. For the SI test and the categorized subjective judgement test, the words composed of the four mora words were explained by the short sentences. The explanation was chosen from “Daijirin[4]: Japanese dictionary”. The presented sound materials are randomly ordered.

2.3 Subjects
Fifteen subjects with normal hearing ability participated in the experiment.

3 Results
3.1 Speech intelligibility test score
The WI test score and the SI test score were plotted versus its speech levels in Figure 2. The scores are the percentage of correct responses obtained from fifteen subjects’ answer for each condition (15 persons *10 words = 150 responses). The relationships between the speech levels and the speech intelligibility scores were modelled for each condition by using logistic regression analysis.

![Figure 2: Speech intelligibility test score with each test condition](image)

- Comparing the intelligibility test scores of the WI test and the SI test, the SI test scores were 20~40 % higher than WI test scores in almost test conditions. This indicated that the sentences of SI test help the subjects understand the meaning of speech sounds.
- In the results of the Case1 (without the background noise), the intelligibility test scores were bigger than 80 % in almost all conditions.
Comparison the intelligibility scores between the Case2 and the Case3, the intelligibility score of the Case 3 tended to be lower than that of the Case 2. On the other hand, when the speech level was low (25 dB), the intelligibility scores were almost 0% in both test conditions. Moreover, when the speech level was the same with the BGN, the intelligibility scores were higher than 80% in all conditions.

### 3.2 Degree of audibility / understanding of speech

Figure 3 shows the condition means of the “degree of audibility of speech” and “degree of understanding of speech” across subjects for the speech levels. The mean values were obtained from fifteen subjects’ answer for each condition (15 responses).

- From the Figure 3, when the speech level increased, the degree of audibility and understanding of speech increased regardless of the three test cases. The subjective judgements of Case 2 were relatively higher than that of Case 3.
When there is no background noise (Figure 3(a)), the degree of audibility and understanding of speech were considerably high, even if the speech levels were low (20 dB).

From Figure 3(a), when the speech levels were low, the degree of understanding of speech was higher than that of audibility. On the other hand, when the speech levels were low (25, 30 dB) the degree of audibility of speech was higher than that of understanding in Figure 3(b) and Figure 3(c). When the speech levels were 35, 40, and 45 dB, the degrees of both were almost the same in all test conditions. From this experimental result, it was suggested that the judgement criteria of the audibility was affected the background noise. Further studies about the impression of speech audibility are necessary.

3.3 Relationships between speech intelligibility score and Degree of audibility / understanding of speech

In order to investigate the relationships between the degree of audibility and understanding on speech and the speech intelligibility test scores, the intelligibility scores with the same category judgement were calculated. The averaged speech intelligibility scores were plotted across the categories (Figure 4).

Figure 4: Relationships between speech intelligibility and subjective judgement
When the degree of audibility and understanding of speech increased, the speech intelligibility test scores also increased. When the degree values were “1. Not at all” and “2. Slightly”, there is no significant difference with the intelligibility test scores regardless of test method.

From the Figure 4, when the degree of audibility of speech was “1. Not at all”, the WI and SI test score were 0 ~ 1 %. In addition, when the degree of that was “2. Slightly”, the WI test scores were about 9 ~ 10 % and the SI test score were 11 ~ 13%.

When the degree of understanding of speech was “1. Not at all”, the WI and the SI test score were 3~5 %. Moreover, when the degree of that was “2. Slightly”, the WI test scores were about 7 ~ 9 % and the SI test score were higher than the WI test (17 ~ 29 %).

4 Conclusions

In order to find the relationships between the speech intelligibility test and the degrees of audibility and understanding of speech, we examined the word intelligibility test and the sentence intelligibility test. In addition, “degree of audibility of speech” and “degree of understanding of speech” were also investigated by two categorized subjective judgement test. We obtained the following results.

1. When the speech levels are 20~45 dB and there is no background noise, the intelligibility test scores are higher than 80 %. Moreover, the degrees of audibility and understanding of speech are considerably high.

2. The degree of audibility and understanding of speech shows a relatively high correlation with the speech intelligibility test score. When the degree of audibility of speech was “1. Not at all”, the WI and SI test score were 0 ~ 1 %. Besides, when the degree of that was “2. Slightly”, the WI test scores were about 9 ~ 10 % and the SI test score were 11 ~ 13%.

3. When the background noise was presented, the degree of audibility was generally higher than that of understanding. On the other hand, the degree of audibility was lowered than that of understanding without the background noise. This is considered the judgement criteria of the speech audibility is affected by the background noise. Further studies about the impression of speech audibility are necessary.

Acknowledgments

The authors wish to thank all the subjects for spending their time to do the test.

References