ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Architectural Acoustics
Session 3aAAb: Architectural Acoustics Potpourri

3aAAb2. Subjective experiment on suitable speech-rate of public address announcement in public spaces
Sakae Yokoyama* and Hideki Tachibana

*Corresponding author's address: Institute of Industrial Science, The University of Tokyo, Komaba 4-6-1, Meguro-ku, 153-8505, Tokyo, Japan, sakae@iis.u-tokyo.ac.jp

In such public spaces as railway stations, airport terminals, shopping arcades, etc. in big city in Japan, it is so noisy and reverberant that the information provided through public address system can not be clear so often. In case of disaster such as a big earthquake, a fire, etc., it is important to ensure the intelligibility of announcement through public address system for evacuation. In order to ensure not only comfortableness but also safety in public spaces, authors have performed field measurement and subjective experiment in a simulated sound field. The experimental results indicated that the difficulty in listening to public address announcement can be caused by surrounding noise, speed of the speech and reverberation. As for the influence of background noise and reverberation on intelligibility, subjective hearing test has systematically performed in a 3-dimensional simulated sound field. In this study, the speeds of public address announcement in actual case were measured and the effect of the speech-rate of announcement in Japanese on the intelligibility was investigated by laboratory experiments. For changing the speed of speech, TTS (Text To Speech) software was applied. The influence of native/non-native language on intelligibility was also investigated by obtaining the cooperation non-native students.

Published by the Acoustical Society of America through the American Institute of Physics
INTRODUCTION

In such public spaces as railway stations, airport terminal buildings and shopping arcades, acoustical problems of uncomfortableness and difficulties in listening to the public address announcement are often caused by poor acoustical treatment. Particularly, in case of disaster such as fire and earthquake, accurate information should be provided as well as ordinary situations. This problem is serious not only for elderly and hearing-handicapped people but also for foreigners. In the design of this kind of spaces, therefore, acoustical consideration is also essential as well as those for safety, fire prevention and maintenance. Recently, some researches have been performed for the elderly and visual/auditory handicapped persons, but those for foreigners have been very few so far.

To investigate such acoustical problems, the authors have made acoustical measurements in various public spaces and a series of laboratory experiments [1-3]. In this study following these studies, to examine the effect of the speech-rate of spoken announcements, speech intelligibility and impression of listening difficulty were investigated by obtaining the cooperation of foreign students studying in Japan as well as Japanese students. To realize the natural aural impression in an acoustic laboratory, the 6-channel recording/reproduction technique developed by the authors was applied [4]. By the use of this technique, the sound fields in an original space can be correctly reproduced with a correct 3-dimensional auditory sensation.

FIELD MEASUREMENTS

To investigate the speech rate of spoken announcements in real typical public spaces, we have performed field surveys in public spaces in big cities in Japan (see Fig. 1). In these surveys performed in busy daytime, environmental sounds including programmed/real-time announcements through public address system were recorded through a sound level meter set at a height of 1.2 m above the floor. At the same time, the environmental sounds were recorded through a 6-channel microphone system [4] for the subjective experiments mentioned below. During the field survey in an airport-terminal building, a large number of flights were canceled or delayed owing to a heavy rain, the real-time announcements by airport staff were also recorded.

Speech Rate

In these field surveys, it was often difficult to listen to the real-time announcements, because of unsuitable speech rate and uncomfortably loud. From the recordings including spoken announcements in railway stations, the speech rates were obtained (see Fig. 2). In these results, it can be seen that the speech rate of real-time announcement about 1.3 times as fast as programmed ones on average; in the real-time announcements, the rate range from 4.9 mora/sec. to 9.5 mora/sec. and in the programmed case, it ranges from 3.6 mora/sec. to 6.5 mora/sec.

Fluctuation

From the recordings for ten minutes, the fluctuation in A-weighted sound level was also obtained. Two examples of the measurement results are shown in Fig. 3. In both cases, it is seen that the programmed announcements under normal condition were set at a suitable level both in Japanese and in English. On the other hand, the real-time announcements by airport/station staff were made at unsuitable level.
FIGURE 1. Public spaces in big cities in Japan.

FIGURE 2. The speech rates of public address announcements in railway stations.

FIGURE 3. The fluctuation in A-weighted S.P.L. of environmental sounds including public address announcements.
SUBJECTIVE EXPERIMENT

In order to investigate the influence of the speech rate of public address announcement on speech intelligibility, the following subjective tests were performed in an artificially simulated sound field using recorded environmental sounds in real public spaces.

Sound Field Simulation

Figure 4 shows the diagram of the system used in the subjective experiments; system A is for the sound field simulation of speech (sentence) through public address system and system B is for the actual environmental sounds. The 6-channel reproduction system is applied to both systems. At the center point of the reproduction system, the sound heard in the real sound field can be reproduced with correct 3-dimensional information. The sound pressure level at the center of the reproduction system (listening position) was adjusted so as to be equal to the level measured in the real sound field. In system A, the directional impulse responses measured by using the 6-channel microphone system consists of six uni-directional microphone units (Sanken, CU-6CH) and dry source were convolved. In system B, the recorded 6-channel signals were mixed with the convolved directional signals mentioned above and reproduced from the six loudspeakers set in an anechoic room (see Fig. 5).

FIGURE 4. 6-channel recording/reproduction system

FIGURE 5. Subjective test in the simulated sound field.
Listening Difficulty

Experimental Procedure

Using the recordings including spoken announcements, the subjective experiment on listening difficulty was performed in an anechoic room by applying the 6-channel recording/reproduction system (see Fig. 4) [4]. As the test sound, the environmental sounds recorded in thirty different public spaces in Japan including the public address announcements in Japanese were used. The time duration of respective test sounds was 15 seconds. The time-averaged A-weighted sound pressure levels ($L_{A,15s}$) ranged from 61.1 dB to 84.0 dB. The speech rate ranged from 4.3 mora/sec. to 8.9 mora/sec. In the simulated sound fields, the test subject sitting at the listening position was asked to image being in a public space and to judge his/her impression for each space item by item (“difficult”, “reverberant”, “fast” and “noisy”) in four-step categories shown in Table 1. As the test subjects, ten Japanese university students in twenties and twelve foreign students in twenties/thirties participated in this experiment.

| TABLE 1. 4-step category common to other items; reverberant, fast, noisy |
|------------------------|-----------------|-----------------|
| 4. Extremely difficult  | “Difficult”     | “Difficult”     |
| 3. Fairly difficult     | 0.421*          | 0.695**         |
| 2. A little difficult   | 0.705**         | 0.643**         |
| 1. Not difficult        | 0.854**         | 0.911**         |

<table>
<thead>
<tr>
<th>TABLE 2. Correlation coefficient between “Listening difficulty” and other items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese students (native)</td>
</tr>
<tr>
<td>item</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>“Reverberant”</td>
</tr>
<tr>
<td>“Fast”</td>
</tr>
<tr>
<td>“Noisy”</td>
</tr>
</tbody>
</table>

* significance level; 0.05, ** significance level; 0.01

Experimental Result

From the results of Japanese students (see Table 2), the score for “listening difficulty” is highly correlated with the score for “noisy” and “fast” ($p<0.01$). It seems that “difficulty in listening to native language” is influenced by “noisiness” of surrounding environmental noise and “speed” of public address announcement. The relationship between the score for “fast” and the speech rate of the announcement was shown in Fig. 6, in which the experimental result is highly correlated with the speech rate. It seems that the score for “speed” of real-time announcement is higher than that of programmed one. In this result, it is seen that the responses of category 2, “a little fast” appeared when the speech rate of announcement was 7.0 mora/sec. In case of programmed one, the speech rate is nearly 5.0 mora/sec.
Speech Intelligibility

Test Sound

For the preparation of the intelligibility tests, 70 Japanese words (CD-ROM; four mora, word familiarity 5.5-7.0) were selected from the NTT data base. A plain and short sentence explaining each of the word in Japanese was composed referring a Japanese dictionary (see Table 3). From these sentences, female phonetic signals (dry sources) were synthesized by using TTS soft “Voice Text”. In order to investigate the influence of the speech rate on speech intelligibility and the impression of listening difficulty, the speech rate of the synthesized sentence was set at from 4.0 mora/sec. to 8.3 mora/sec. in five steps based on the results of the field measurements as a reference. For the intelligibility tests, the directional impulse responses were measured through the 6-channel receiving system in an extremely reverberant gymnasium with a room volume of about 2,000 m³ and reverberation time of 5.0 sec. in 500 Hz octave band. In the acoustical measurement, the receiving system was set at a distance of 10 meters from a dodecahedral loudspeaker system. In order to examine the influence of the reverberation on suitable speech rate, the reverberation time of each directional impulse response was artificially changed by applying signal processing technique. Experimental condition of reverberation time in 500 Hz octave band was set at three steps of 1.0, 2.0, 3.0 sec., as shown in Fig. 7. The synthesized sentences were convolved with each directional impulse response and reproduced through the six loudspeakers. The reproduction level in time-averaged $L_A$ of each sentence was set at 65 dB. As the back ground noise, an air conditioning noise recorded by 6-channel receiving system with 55 dB in A-weighted sound level ($L_A$) was used.

### TABLE 3. Examples of sentences (in Japanese)

<table>
<thead>
<tr>
<th>Sentence (Explanation in Japanese)</th>
<th>Answer (Word in four mora); in Hiragana or Katakana or Romaji (roman letters)</th>
<th>Meaning in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>その年、初めて降る雪のことを初雪という</td>
<td>はつゆき／ハツユキ／Hatsuyuki</td>
<td>First snow of the season.</td>
</tr>
<tr>
<td>不快な感じを起こさせる音のことを雑音という</td>
<td>ざつおん／ザツオン／Zatsuon</td>
<td>Noise</td>
</tr>
<tr>
<td>よく調べ考えて真理を極めることを研究という</td>
<td>けんきゅう／ケンキュウ／Kennkyu</td>
<td>Study</td>
</tr>
</tbody>
</table>

![FIGURE 7. Waveforms of impulse responses artificially modified.](image)

Experimental Procedure

In the hearing test, four sentences were presented every five second interval in each experimental condition. Each subject listened to a set of 60 test signals in total (4 sentences, 15 test conditions). The subject was instructed to write down the word explained in each sentence which he/she could catch in “kanji” (Chinese character), “hiragana/katakana” (Japanese syllabary) or “romaji” (roman letters) and simultaneously rate the listening difficulty into one of the four categories as shown in Table 1. Before the listening test, each test subject listened to eight sentences (two conditions) as an exercise.
Test Subjects

In this experiment, ten Japanese students and nine foreign students in their twenties and thirties participated as the test subjects. The foreign students, four from Korea, four from Taiwan and one from Thailand, are studying in Japan and are good at everyday conversation in Japanese. The length of their stay in Japan is from two months to six years.

Experimental Result

The intelligibility score was calculated as the percentage of the test signals correctly written down. The average score were obtained for respective test conditions in each subject group (native/non-native). Although the test on the impression of listening difficulty was performed using ordinal scale as mentioned above, the responses were treated as in interval scale and the average was calculated.

Figure 8 and Figure 9 show the relationship between the experimental result (intelligibility score/Listening difficulty) and the value of speech rate of each test condition. In Figure 8, the difference between the two subject groups (native/non-native) is clearly seen. For the native subjects, the intelligibility score are higher than 80 % except for the condition where the speech rate was 8.3 mora/sec. and the reverberation time was 3.0 sec. It indicates that the speech rate and the reverberation do not have serious influence on the intelligibility for native language. On the other hand, the score for the non-native subjects are clearly lower compared with that for the native subjects. From these results, it is seen that the reverberation has a serious influence on speech intelligibility for non-native language, although the influence of the speech rate is not clearly seen. Especially in the case where the reverberation time was 3.0 sec., the score for non-native subjects was about 50 %. When considering this result from a viewpoint of safety, this is a serious problem for public spaces in which the reverberation time is longer than 3.0 sec. In Figure 9, in both conditions where the reverberation time was 2.0 sec. and 3.0 sec., the difference between the two subject groups (native/non-native) is clearly seen. From the results, the tendency is seen that the listening difficulty judged by the native subjects is influenced by speech rate and reverberation, although their influences on the speech intelligibility score are not so serious. On the other hand, the listening difficulty for the non-native subjects was much more influenced by speech rate and reverberation conditions. In the case where the reverberation time was 3.0 sec., the response fell into category 3 “fairly difficult” or category 4 “extremely difficult”.

FIGURE 8. Experimental results of intelligibility.  FIGURE 9. Experimental results of listening difficulty.
CONCLUSIONS

In order to investigate the transmission characteristics of speech announcement in public spaces, two steps of subjective experiments were performed in this study by using the 6-channel recording/reproduction system for sound field simulation.

Firstly, the primary factor of listening difficulty of the P.A. announcement was examined. As a result, it has been found that the listening difficulty is related to the surrounding noise level, the speech rate of announcements and the reverberation.

Next, the effect of speech rate on the speech intelligibility and the impression of listening difficulty was examined by artificially synthesized sound field simulation, in which the speech rate and the reverberation time were changed in steps. In this experiment, in order to investigate the influence of the background of language which is his/her mother’s tongue or not, foreign students cooperated on the experiment. As a result, it has been indicated that the speech rate and the reverberation do not have serious influence on the speech intelligibility for native language, whereas the reverberation is a serious factor for the speech intelligibility for non-native language. As for the impression of the listening difficulty, it has been indicated that both of the speech rate and the reverberation are influential in both cases for native/non-native language. Especially when considering emergency cases, the listening difficulty for non-native language could be reduced by keeping suitable speech rate of the public address announcements.

ACKNOWLEDGMENTS

This study was carried out as a part of Grants-in-aid for Scientific Research (Basic research program (A) 22241040).

REFERENCES