Effect and Sound Localization with Dichotic-Listening Digital Hearing Aids

Atsunobu Murase1,2,*, Fumie Nakajima1, Shuichi Sakamoto1, Yōiti Suzuki1, and Tetsuaki Kawase3

1Research Institute of Electrical Communication and Graduate School of Information Sciences, Tohoku University
2Matsushita Electric Industrial Co., Ltd.
3Departments of Otolaryngology, Tohoku University Graduate School of Medicine

murase@ais.riece.tohoku.ac.jp

Abstract

Sensorineural hearing loss is commonly accompanied by reduced frequency selectivity, which causes the larger adverse effect of masking. To cope with this issue, some studies have addressed the effects of dichotic listening to reduce the masking between contiguous frequency bands. This paper investigates sound image localization perceived in both dichotic and diotic listening conditions, as well as speech intelligibility. The first experiment carried out a speech intelligibility test using vowel-consonant-vowel nonsense syllables presented to four subjects with mild to severe sensorineural hearing loss. Speech stimuli were divided into two frequency bands for dichotic listening to reduce masking to a consonant by the preceding vowel. In consideration of formant frequencies of Japanese vowels, the frequency boundary was set to 0.8 kHz (between F1 and F2) and 1.6 kHz (middle of F2). Results showed that greatest improvement in intelligibility was observed when the dividing frequency was 0.8 kHz. The second experiment conducted sound image localization using daily conversation speech samples for six young adults with normal hearing and three elderly persons with hearing impairment. Head-related transfer functions (five directions, non-individualized) were convolved to realize directivity of the speech samples. Results show that localization errors varied with the dividing conditions and whether hearing was impaired or not.

1. Introduction

Sensorineural hearing loss is typically characterized by increased hearing thresholds, reduced dynamic range of hearing, degraded of temporal resolution and increase in temporal masking, and reduced frequency selectivity. Among them, the reduction of frequency selectivity engenders remarkable disadvantages through large and extensive masking, particularly masking of middle and high frequency components by intense low frequency components: the so-called upward spread of masking.

Several researchers have examined effects of dichotic listening to cope with the upward spread of masking, [1–6]. Dichotic listening originally means listening to a different signal in each ear, but in this context it means listening to differently filtered speech sound in each ear. That is, the speech signal is divided into two complimentary parts in terms of frequency spectra to reduce masking between contiguous frequency bands. Previous studies have suggested that dichotic listening is basically effective to improve speech intelligibility [1–2, 4–6], but one report cited ineffectiveness in noisy environments [3]. Nevertheless, many questions remain in these studies. What the best bandwidth and deviation, e.g. partition frequency, particularly in a noisy environment? Moreover, sound image localization under a dichotic listening condition remains an important problem.

This paper addresses two aspects of the effect of dichotic listening. One topic is the intelligibility test results from four hearing-impaired listeners. Sound stimuli were divided into two frequency bands to largely reduce the masking of a consonant by a preceding vowel. The other topic is the perceived direction of speech in dichotic listening. Head-related transfer functions were convolved to realize speech sample directivity. We studied six young adults with normal hearing and three hearing-impaired elderly persons.

2. Speech intelligibility test

2.1. Experimental procedure

The listening test was conducted in a soundproof room. Speech stimuli were presented to listeners through headphones (HDA-200; Sennheiser Electronic GmbH & Co. KG). Speech stimuli were nonsense Vowel-Consonant-Vowel (VCV) syllables uttered by a native Japanese female talker. The first vowels in each
VCV syllable were /u/ and followed by one of 67 kinds of CV syllables. For dichotic patterns, the speech signal was split into two bands by a low-pass filter (LPF) and a high-pass filter (HPF). Considering formant frequencies of Japanese vowels, the boundary frequency was set to 0.8 kHz (between F1 and F2) and 1.6 kHz (middle of F2). As the diotic listening condition, all-pass filters (APF) were used instead of LPF and HPF. Table 1 summarizes the dividing patterns. For dichotic listening, stimuli with amplitude of −6 dB (Dichotic−6 dB) were also prepared because, in terms of binaural summation of loudness, loudness in a dichotic condition is estimated to be about 6 dB lower than in a diotic condition under a moderate sound level [7]. Two kinds of noise, speech spectral noise and road noise recorded on highway (expected strong upward spread of masking), were added to the speech signal at signal-to-noise ratios (S/N) of 4 and 0 dB in terms of the A-weighted level (Table 2). The level of the speech signals was kept at the listener’s most comfortable level (MCL), which was determined whereas the speech signal was presented to both ears without noise. The interval between stimuli was set to 3 s. Listeners were four elderly persons with mild to severe sensorineural hearing losses. Table 3 shows hearing threshold levels and most comfortable levels (MCL) for each listener. Listeners were asked to write the perceived syllable.

2.2. Result

Intelligibility scores for listener A and averages of all listeners are shown in Figs. 4 and 5, respectively. Figure 4 shows that for Listener A, the intelligibility score with a Dichotic0.8 pattern is about 15% higher than that in a quiet condition; scores decrease according to the worsening S/N. With a Dichotic1.6 pattern, scores are generally lower than those with other patterns. Average scores in Fig. 5 show a similar pattern. We examined results with a two-way repeated-measure ANOVA for more detailed analysis. In this analysis, dividing patterns and noise conditions were treated as between-subject variables and listeners were treated as a repeated-measure. Dividing patterns and noise conditions were inferred to be statistically significant at $p < 0.01$. Lists arranged in descending order of the mean value are the following:

Dichotic0.8 > Diotic > Diotic−6dB > Dichotic1.6

multiple comparison tests showed significant differences among the following conditions:

- Dichotic0.8 and Diotic−6dB
- Dichotic0.8 and Dichotic1.6
- Diotic and Diotic−6dB
- Diotic and Dichotic1.6

<table>
<thead>
<tr>
<th>pattern</th>
<th>left</th>
<th>right</th>
<th>cross over</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diotic</td>
<td>APF</td>
<td>APF</td>
<td>–</td>
</tr>
<tr>
<td>Dichotic0.8</td>
<td>LPF</td>
<td>HPF</td>
<td>0.8</td>
</tr>
<tr>
<td>Dichotic1.6</td>
<td>LPF</td>
<td>HPF</td>
<td>1.6</td>
</tr>
<tr>
<td>Diotic−6dB</td>
<td>APF</td>
<td>APF</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 1: Frequency dividing patterns

<table>
<thead>
<tr>
<th>condition</th>
<th>kind of noise</th>
<th>S/N[db] ($L_{Aeq}$)</th>
<th>S/N[db] ($L_{eq}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>quiet</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>speech 4dB</td>
<td>speech spectral noise</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>speech 0dB</td>
<td>speech spectral noise</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>road 4dB</td>
<td>road traffic noise on highway</td>
<td>4.0</td>
<td>-12.7</td>
</tr>
<tr>
<td>road 0dB</td>
<td>road traffic noise on highway</td>
<td>0.0</td>
<td>-16.7</td>
</tr>
</tbody>
</table>

Table 2: Noise condition

<table>
<thead>
<tr>
<th>Subject</th>
<th>Hearing threshold level [dB]</th>
<th>MCL [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left ear</td>
<td>Right ear</td>
</tr>
<tr>
<td>Listener A</td>
<td>37.5</td>
<td>32.5</td>
</tr>
<tr>
<td>Listener B</td>
<td>53.7</td>
<td>52.5</td>
</tr>
<tr>
<td>Listener C</td>
<td>41.3</td>
<td>46.2</td>
</tr>
<tr>
<td>Listener D</td>
<td>61.2</td>
<td>71.3</td>
</tr>
</tbody>
</table>

Table 3: Hearing threshold levels and most comfortable levels on intelligibility tests for each listener
3.3. Result

Judgment results (Diotic, Dichotic0.8L and Dichotic1.6L) by all the normal hearing listeners and by hearing impaired listener B are shown in Fig. 6 and Fig. 7, respectively, in which the square measure of each circle indicates the relative number of responses. Comparing Figs. 6 and 7 shows that this hearing impaired listener tends to localize sound stimuli to the side from which the low frequency part was presented with Dichotic0.8L and Dichotic1.6L pattern. These results are similar to those for other hearing impaired listeners. In contrast, normal-hearing listeners tend to localize to both directions with Dichotic0.8L and Dichotic1.6L patterns (Fig. 7).

4. Discussion

Speech signals are divided into two bands in intelligibility tests. Here, the dividing frequencies were selected by considering the formant frequencies of Japanese vowels. Results show that simple dichotic presentation with two bands is effective to improve intelligibility for hearing impaired listeners.

However, the effects seem to relatively small in noisy conditions, as observed in previous studies [5].
The sound-image localization experiment showed that hearing impaired listeners tend to localize to the side from which the low frequency portion was presented. This might be attributable to the audiogram of hearing impaired listeners: high frequency hearing loss is more severe than low frequency loss.

5. Conclusions and future work

We examined the effect of dichotic listening to reduce upward spread of masking. In a dichotic listening condition, speech signals were divided into two frequency bands by considering formant frequencies of Japanese vowels. Moreover, we compared sound localization in dichotic and diotic conditions. The following conclusions were thereby obtained:

- Speech intelligibility is improved for hearing-impaired listeners, particularly in quiet conditions and for the dividing frequency of 0.8 kHz;
- For hearing impaired listeners, the sound image tends to locate on the side in which the low frequency portion exists; and
- Normal hearing listeners tend to perceive two sound images which correspond to the low frequency part and the high frequency part.

6. Acknowledgements

This research was partially supported by JSPS, Grant-in-Aid for Scientific Research (B), 14350193, 2002.

7. References