Relationship between broadening of auditory filter bandwidth near 2 kHz and speech discrimination/identification of monosyllables /de/ and /ge/ by young and elderly Japanese listeners

Keiichi Yasu1,*, Takayuki Arai1, Tatsunori Endoh1, Kei Kobayashi2 and Mitsuko Shindo3

1Faculty of Science and Technology, Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo, 102–8554 Japan
2Section of Audiology, School of Population Health, The University of Auckland, 261 Morrin Road, Glen Innes, Auckland, New Zealand
3Research Center for Communication Disorders, Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo, 102–8554 Japan

(Received 12 May 2014, Accepted for publication 23 June 2014)

Keywords: Auditory filter, Speech discrimination, Speech identification, Just noticeable difference, Categorical perception

PACS number: 43.71.Es, 43.71.Lz, 43.71.Ky, 43.72.Qr [doi:10.1250/ast.36.35]

1. Introduction

Various studies have been carried out on the relationship between the auditory filter [1,2] and speech perception. It is reported that the auditory filter bandwidth of a listener with hearing loss was broader than that of a listener with normal hearing, and frequency selectivity was reduced owing to this broadening [3]. Thibodeau and Van Tasell investigated the relationship between the equivalent rectangular bandwidth (ERB) [3] of the auditory filter at 2 kHz and the identification of the monosyllables /de/–/ge/. As a result, speech identification performance was reduced when the auditory filter was broad at 2 kHz [4]. They concluded that the resolution near 2 kHz was important because the second formant (F2) transition existed near 2 kHz when the stimuli included the vowel /i/ [4].

On the other hand, various studies have focused on the discrimination in F2. Lee et al. found that the just noticeable difference (JND) of F2 increased when the performance of speech discrimination was reduced [5]. Abe et al. conducted an experiment on the identification of stimuli in a continuum from /ba/ to /wa/. They found that the slope of the categorical perception decreased when the JND of the F2 transition increased; listeners who showed poor performance in the discrimination of the F2 transition had a more gradual slope of /ba/–/wa/ categorical perception [6].

The performance in the discrimination of the F2 transition is important for categorical perception. However, there have been few studies on the relationship between the discrimination of the F2 transition and the auditory filter bandwidth. Therefore, Endoh et al. investigated this relationship for young Japanese listeners with normal hearing [7]. They confirmed by a spectral smearing technique [8] that the JND of the F2 transition increases when the auditory filter bandwidth broadens. Additionally, they reported that the slope of /de/–/ge/ categorical perception decreased when the JND of the F2 transition increases. In this study, we conducted similar experiments on elderly Japanese listeners, and investigated the relationship between the broadening of the auditory filter bandwidth, the discrimination of the F2 transition, and the identification of the monosyllables /de/ and /ge/.

2. Methods

2.1. Participants

Twenty elderly listeners (two males and 18 females; age range, 63 to 80 years; mean age, 72.2 years) and nine young listeners (six males and three females; age range, 22 to 25 years; mean age, 24.5 years) participated in the experiment. All participants were native speakers of Japanese. The participants listened to stimuli presented to either the right or left ear with the low pure-tone threshold at 2 kHz.

2.2. Measurement of pure tone threshold and auditory filter bandwidth

Before the experiments, the pure tone thresholds for both ears were measured using an audiometer (AA-79S, RION) at 2 kHz. The auditory filter bandwidth at 2 kHz was also measured using a measurement system of auditory filters (HD-AF, RION) [9,10]. In the measurement of auditory filters, the presentation level was set at 30 dB in sensation level. Participants listened to the stimuli presented to the right or left ear at pure tone threshold levels.

2.3. Stimuli and procedure

Stimuli were synthesized using a Klatt speech synthesizer [11]. Each stimulus consisted of the formant transition and steady-state parts of a vowel. F0 was set to 100 Hz. F1, F2, F3, F4, and F5 in the steady-state part were set to 450, 1,900, 2,400, 3,500, and 4,500 Hz, respectively. Part of the formant transition was placed at 0–30 ms, and the F1 and F2 frequencies were linearly varied during the period, while the other formants were kept constant. Furthermore, the onset

*Corresponding author. E-mail: k-yasu@ieee.org

© 2015 The Acoustical Society of Japan

Acoust. Sci. & Tech. 36, 1 (2015)
frequency of F1 was fixed at 180 Hz, whereas the onset frequency of F2 was varied. Details of the F2 transition are outlined in the following sections. The duration of the steady-state part was 270 ms, and the total duration of each stimulus was 300 ms. The other parameters were set to the default values in the Klatt speech synthesizer [11].

The presentation level was 30 dB in sensation level, which is the same as in the measurement of the auditory bandwidth. All experiments were conducted in a sound-treated room (AT-81, RION), for approximately 60 minutes including periods of breaks in between experimental sessions.

2.4. Experiment 1: Discrimination of F2 transition

We conducted a discrimination test using pairs of consonant-vowel (CV) monosyllabic stimuli; within each pair, the F2 transitions were increased and decreased. A three-interval three-alternative forced-choice task was used (3I-3AFC). In each trial, there were three stimuli with each of them taken from one of the pairs of CV stimuli. The participants were instructed to choose the stimulus that was different from the other two stimuli. In the first trial, the onset frequencies of F2 were 1,500 and 2,500 Hz, corresponding to the monosyllables /de/ and /ge/, respectively. The onset frequencies of F2 were changed toward 2,000 Hz in a two-up one-down procedure [12], in which the step size was 50 Hz in the first 20 trials and reduced to 20 Hz afterwards. In each 3I-3AFC trial, the onset frequencies of two stimuli were set to 2000 – ΔF Hz, and that of the other stimulus was set to 2000 + ΔF Hz. The two-up one-down procedure was used to track 70% of the psychometric function for measuring the JND of the F2 transition. The JND of the F2 transition was defined as the difference in the onset frequency, 2ΔF Hz.

2.5. Experiment 2: Identification of /de/ vs. /ge/

Listeners identified stimuli from a continuum as /de/ or /ge/. In the continuum, the F2 onset frequency was varied from 1,500 to 2,500 Hz in 50 Hz steps (1,500, 1,550, . . ., 2,500 Hz). There were 21 stimuli in this experiment. During the practice session, to get used to the experimental procedure and stimuli, two monosyllables /de/ (F2 onset frequency: 1,500 Hz) and /ge/ (F2 onset frequency: 2,500 Hz) were presented as stimuli five times. In the main session, all 21 stimuli were presented in random order and repeated ten times. For each trial, after presenting a stimulus, the participants chose the button “/de/” or “/ge/” on a touch panel display.

3. Results and discussion

3.1. Pure tone threshold and auditory filter bandwidth

Table 1 shows the average pure-tone threshold of hearing and the ERB at 2 kHz. There were significant differences between the young and elderly listeners in the pure-tone threshold and ERB. In this study, the average ERB of young listeners at 2 kHz (ERB2) was 374 Hz ± 27.7 Hz and broader than that in a previous study (Glasberg and Moore’s result was 240.7 Hz [3]).

3.2. Experiment 1

The JND of the F2 transition was obtained for each listener as the difference in F2 onset frequency. Table 1 shows the average and standard deviation of JNDS among all listeners. Elderly listeners show a significantly larger JND of the F2 transition than young listeners in the Mann-Whitney U test (p < 0.05). Figure 1(a) shows a scatter plot of the JND of the F2 transition (JND, vertical axis in this figure) versus the auditory filter bandwidth (ERB, horizontal axis in this figure). In Fig. 1(a), a weak correlation was found, that is, the JND of the F2 transition was larger when the ERB was broader. There was a significant correlation between the JND and ERB for all listeners including young and elderly listeners (r = 0.430, p < 0.05). This result suggests that the frequency selectivity near 2 kHz is associated with the JND of the F2 transition.

Table 2 shows the correlation coefficients between the pure-tone threshold at 2 kHz (Threshold), the auditory filter bandwidth (ERB), the JND of the F2 transition (JND), and the slope of categorical perception (Slope) for elderly and young listeners. From this table, the correlation was not statistically significant between the pure-tone threshold and the JND of the F2 transition, which was similar to the result obtained by Abe et al. [6]. Thus, the ERB can predict the JND of the F2 transition better than the pure-tone threshold.

3.3. Experiment 2

An example of the result for one of the listeners is shown as dots in Fig. 2. Each dot is the average value of the /ge/ responses over the ten repetitions for each stimulus (the rate of /ge/ responses) for each F2 transition condition (horizontal axis) for a single listener. The slope of categorical perception [6], i.e., the parameter a, were calculated by fitting with a sigmoidal function. The slope of categorical perception was defined as the slope at the /de/–/ge/ phoneme boundary of the sigmoidal fitting curve where the /ge/ response rate was 0.5. When the listeners’ perception is categorical, the slope of the sigmoidal fitting curve becomes steep at the phoneme boundary [6] (Fig. 2). Figure 1(b) shows a scatter plot of the decreasing slope of categorical perception (slope, vertical axis in this figure) versus the auditory filter bandwidth (ERB, horizontal axis in this figure). A weak correlation was found, that is, the slope of categorical perception decreased when the ERB of the participant’s auditory filter was broad (r = −0.509, p < 0.01). This result suggests that the frequency selectivity at 2 kHz was associated with the identification of /d/ and /g/ categorically [4]. Table 1 also shows the average and standard deviation of the slopes of categorical perception, as well as the significance probabilities in the Mann-Whitney U test.

Table 1 Average and standard deviation of pure-tone thresholds of hearing at 2 kHz (Threshold), auditory filter bandwidth at 2 kHz (ERB), JNDS of F2 transition (JND), and slopes of categorical perception (Slope) for young and elderly listeners. Significance probabilities in the Mann-Whitney U test are also indicated for the comparison between young and elderly listeners.

|                | Elderly   | Young     | U test |
|----------------|-----------|-----------|--------|
| Threshold (dB) | 23.5 ± 2.00 | 6.11 ± 2.54 | (p < 0.05) |
| ERB (Hz)       | 532 ± 43.3 | 374 ± 27.7 | (p < 0.05) |
| JND (Hz)       | 703 ± 41.0 | 564 ± 39.0 | (p < 0.05) |
| Slope          | 0.11 ± 0.00 | 0.16 ± 0.09 | (p < 0.1) |

Acoust. Sci. & Tech. 36, 1 (2015)
Figure 1(c) shows a scatter plot of the slope of categorical perception (slope, vertical axis in this figure) versus the JND of the F2 transition (JND, horizontal axis in this figure). A weak correlation was found, that is, the JND of the F2 transition was larger when the slope of categorical perception decreased ($r = -0.446$, $p < 0.05$). This result suggests that both JND and the slope of categorical perception were associated with increasing ERB.

### 4. Conclusion

In this study, the relationship between the discrimination/identification of F2 and the auditory filter bandwidth near 2 kHz was investigated for young and elderly listeners. As a result, the ERB of the auditory filter of elderly listeners at 2 kHz was broader than that of young listeners. When the auditory filter bandwidth was broad, the JND of F2 transition was larger, and the slope of categorical perception decreased. These results suggest that the auditory filter bandwidth was associated with the JND of the F2 transition and the slope of categorical perception.

**Table 2** Correlation coefficients of pure-tone threshold of hearing at 2 kHz (Threshold), auditory filter bandwidth (ERB), JNDS of F2 transition (JND), and slopes of categorical perception (Slope) for all, elderly, and young participants. The entries marked by * indicate a significant condition at 1%, marked by ** indicate that at 5%.

|        | ERB  | JND  | Slope  |
|--------|------|------|--------|
| All    | 0.659* | 0.257 | -0.183 |
|        | Threshold |      |        |
|        | ERB  | 0.430* | -0.509** |
|        | JND  | 0.446* |        |
| Elderly| 0.586** | 0.070 | -0.137 |
|        | Threshold |      |        |
|        | ERB  | 0.366  | -0.618** |
|        | JND  | -0.412 |        |
| Young  | 0.371 | -0.279 | 0.534 |
|        | Threshold |      |        |
|        | ERB  | 0.012  | -0.111 |
|        | JND  | -0.383 |        |

Figure 1(c) shows a scatter plot of the slope of categorical perception (slope, vertical axis in this figure) versus the JND of the F2 transition (JND, horizontal axis in this figure). A weak correlation was found, that is, the JND of the F2 transition was larger when the slope of categorical perception decreased ($r = -0.446$, $p < 0.05$). This result suggests that both JND and the slope of categorical perception were associated with increasing ERB.

**Fig. 2** The dots shows an example of the result for one of the young listeners in Experiment 2. The solid line is the sigmoidal fitting curve calculated from the results. The dashed line is the slope of the categorical perception at the phoneme boundary where the response rate was 0.5.

### Acknowledgment

This research was partly supported by Sophia University Open Research Center from MEXT.

### References

[1] H. Fletcher, “Auditory patterns,” *Rev. Mod. Phys.*, 12, 47–61 (1940).
[2] E. Zwicker, “Subdivision of the audible frequency range into critical bands (Frequenzgruppen),” *J. Acoust. Soc. Am.*, 33, 248 (1961).
[3] B. R. Glasberg and B. C. J. Moore, “Auditory filter shapes in subjects with unilateral and bilateral cochlea impairment,” *J. Acoust. Soc. Am.*, 79, 1020–1033 (1986).
[4] L. M. Thibodeau and D. J. Van Tasell, “Tone detection and synthetic speech discrimination in band-reject noise by hearing-impaired listeners,” *J. Acoust. Soc. Am.*, 82, 864–873.
(1987).
[5] S. H. Lee, T. Yoshino and K. Ashihara, “Japanese vowel formant frequency discrimination by people with hearing impairments: Position of the second formant,” Jpn. J. Spec. Educ., 36(3), pp. 1–9 (1998) (in Japanese).
[6] M. Abe, S. Saito and K. Endo, “Japanese /ba/wa/ identification and formant transition discrimination: Persons over sixty years of age compared to younger adults,” Jpn. J. Spec. Educ., 40(1), pp. 13–23 (2002) (in Japanese).
[7] T. Endoh, T. Arai, K. Yasu, K. Kobayashi and M. Shindo, “Relationship between spread of auditory filter bandwidth and speech discrimination/identification: Focus on the second formant near 2 kHz,” Proc. Autumn Meet. Acoust. Soc. Jpn., pp. 50–51 (2010) (in Japanese).
[8] T. Baer and B. C. J. Moore, “Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech,” J. Acoust. Soc. Am., 94, 1229–1241 (1993).
[9] T. Nakaichi, K. Watanuki and S. Sakamoto, “A simplified measurement method of auditory filters for hearing-impaired,” Acoust. Sci. & Tech., 24, 365–375 (2003).
[10] T. Nakaichi, K. Watanuki and S. Sakamoto, “Development of a measurement system of auditory filters for evaluating the auditory characteristics of hearing-impaired listeners,” J. Acoust. Soc. Jpn. (J), 62, 214–223 (2006) (in Japanese).
[11] D. H. Klatt and L. C. Klatt, “Analysis, synthesis, and perception of voice quality variations among female and male talkers,” J. Acoust. Soc. Am., 87, 820–857 (1990).
[12] H. Levitt, “Transformed up-down methods in psychoacoustics,” J. Acoust. Soc. Am., 46, 467–477 (1971).