Perception of /r/ and /l/ in quiet and multi-speaker babble noise by Japanese and English native listeners *

◎Hinako Masuda and Takayuki Arai
(Faculty of Science and Technology, Sophia University)

1 Introduction

Speech perception in noise is difficult for all listeners. Non-native listeners' speech perception is undoubtedly more difficult compared to that of the native listeners [1-3]. Previous studies have found that even advanced level learners' or bilinguals' perception of foreign sounds under noisy listening conditions fall short of that of the native listeners' performance, even if their performance is similar to the native listeners under quiet listening condition [4, 5].

Specifically in the case of Japanese native listeners, the perception of the English sounds /r/ and /l/ is often viewed as one of the major difficulties the learners face. Perceptual experiment by Adachi *et al.* [6] found that Japanese native listeners were able to identify /r/ and /l/ at approximately 65% under quiet listening condition, but declined to approximately 55% under noise (SNR = －15 dB) while American English native listeners were able to identify them 100% in quiet.

Similarly in Ueda *et al.*'s [7] experiment on the identification of /r/ and /l/, they found that Japanese native listeners' identification rate was approximately 70% under quiet condition, but declined to approximately 55% in noise (SNR = －21 dB). American native listeners' performance, on the other hand, was 100% in quiet and 70% in noise. Their study also suggested that training improves the perception of foreign sounds – the non-native listeners' performance improved in noise after going through training sessions under quiet listening condition.

In the experiment conducted by Akahane-Yamada *et al*. [8], Japanese native listeners underwent 150 to 200 minutes of training sessions, which consisted of the identification of English /r/ and /l/ minimal pairs. Their results suggested that perceptual training is effective in improving perception of foreign sounds, as well as for the improvement of production. Furthermore, the training effect was maintained even after three to six months.

Studies concerning non-native listeners with advanced level proficiency, Rogers *et al.* [4] compared the perceptual ability by American English native listeners and Spanish-English bilingual listeners who were exposed to English before age six. Although they found a significant difference between native and bilingual listeners in perceiving English monosyllabic words embedded in noise and reverberation, the two groups' performances were similar in quiet listening condition.

Additionally, Mayo *et al.* [5] examined the perception of monosyllabic words with high and low predictability by English-Spanish bilingual listeners, and found that early exposure to a second language is advantageous in perceiving second language sounds in noise. However, they reported that even bilinguals who had been exposed to a second language since infancy did not reach native-level performance.

In sum, second language perception in noisy and reverberant listening conditions are difficult for non-native listeners, but the perception in quiet listening condition may reach native-like level, especially for advanced level learners. However, the tendencies seem to vary with the non-native population. The focus of the present study is to examine how second language proficiency affects perception of /r/ and /l/ in both quiet and noisy listening conditions with varying signal-to-noise ratios. The aims of the experiment

* Perception of /r/ and /l/ in quiet and multi-speaker babble noise by intermediate and advanced Japanese learners of English and English native listeners, by MASUDA, Hinako and ARAI, Takayuki (Sophia University).

are: 1) to examine the differences in performance between Japanese native listeners with intermediate and advanced level English proficiency, and 2) to examine the difference in performance between advanced level Japanese native listeners and English native listeners.

2 Perceptual experiment

2.1 Participants

Thirty listeners participated in the present experiment, and they are divided into three groups according to their language learning background (see Table 1). Three Japanese participants were excluded from analyses due to insufficient participant information. None of the participants reported any hearing problems.

Table 1 Data of participants

	Intermediate learners	Advanced learners	English natives
Number of participants	N = 8	N = 13	N = 6
Age (Range)	Mean 23.0 (20-31)	Mean 26.4 (20-35)	Mean 20.8 (20-21)

Twenty-four Japanese native listeners participated as the non-native listener group (15 male, 9 female), and six English native listeners (2 male, 4 female) participated as the control group. Japanese native listeners are further divided into two groups according to their English proficiency levels. Participants with higher proficiency were grouped as advanced level learners of English, who had achieved higher than 850 on TOEIC® examination [9] or achieved equivalent scores on TOEFL® examination [10], and/or were placed in advanced level English classes at a university in Japan.

The remaining participants are grouped as intermediate level learners of English, who had achieved below 650 on TOEIC® examination, and/or were placed in intermediate level English class at a university in Japan. Participants who do not have experience of living abroad received English education from age twelve at junior high schools and high schools in Japan. None of the participants reported any hearing problems.

Six English native listeners participated in the experiment as control listeners. All six participants were recruited in Japan. All had considered themselves as native speakers of American English, although three participants had Japanese, Chinese, or Korean background who were either born and raised in the United States or moved to the United States in the first few years of life and were exposed to American English thereafter. Those three participants considered English as their native language, and Japanese, Chinese, or Korean as a foreign language. The level of Japanese was in the range of intermediate in all participants. All English native listeners had less than one year of residence in Japan.

2.2 Stimuli

Twenty-three English consonants /b, tʃ, d, f, g, h, dʒ, ʒ, k, l, m, n, p, ɹ, s, ʃ, t, θ, ð, v, w, j, z/ were embedded in /ɑ __ ɑ/ context and presented to the listeners. The data of /r/ and /l/ are analyzed in this paper. The speaker of the stimuli is a female Japanese-English bilingual speaker. The stimuli were recorded in a sound-proof room, using a digital sound recorder (Marantz PMD 660) and a microphone (SONY ECM-23F5) at a sampling frequency of 48 kHz. The stimuli were later downsampled to 16 kHz.

Participants first listened to the stimuli with background noise (multi-speaker babble noise at SNR = 0 dB, 5 dB, 10 dB), and then proceeded to the stimuli in quiet. The order of the SNR in the noisy conditions was randomized. The stimuli in noise were preceded and followed by 1 second of noise. The added multi-speaker babble noise was taken from NOISEX [11]. Multi-speaker babble noise was selected as background noise because it resembles real-life environment that second language learners may experience difficulties in foreign language perception, such as in a school cafeteria.

2.3 Procedure

A laptop computer was used to present the stimuli and to record the listeners' responses. All experimental procedures were conducted by using Praat [12]. Participants were presented with the stimuli through USB audio amplifier (ONKYO MA-500U) and headphones (STAX SR-303 and STAX SRM-323A). The laptop computer and audio amplifier were digitally connected via USB interface.

All participants were given 23 practice trials (six each in SNR = 0 dB, 5 dB, 10 dB, and five in quiet). The practice trials were not scored. After the practice trials, participants proceeded to the main experiment where 460 trials were presented (345 in multi-speaker babble noise and 115 in quiet). They were asked to listen to each stimulus and to choose the best consonant that fits to what they heard from the table of 23 consonants, as shown in Figure 1.

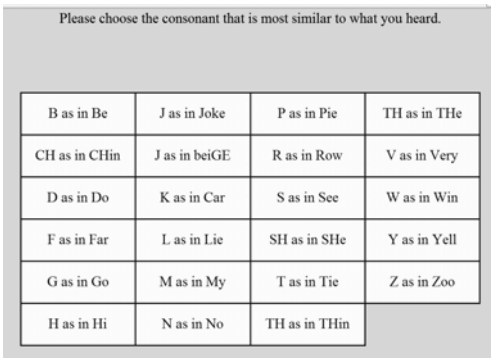

Figure 1 Experimental interface
(Reference: Cutler *et al.*, 2004)

3 Results

The general results of the three listeners' perception of /r/ and /l/ are shown in Figures 2 and 3, respectively.

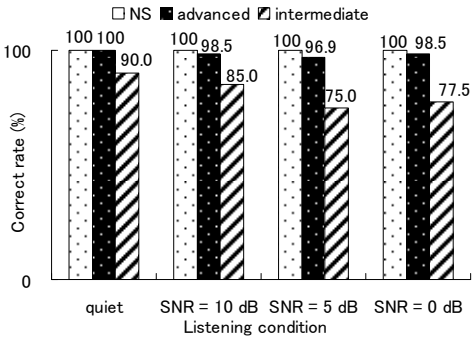

Figure 2 Correct rates of the perception of /r/ by the three listener groups

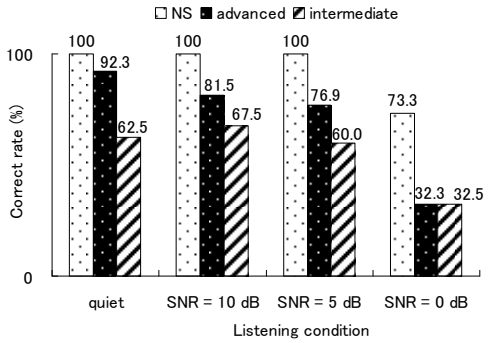

Figure 3 Correct rates of the perception of /l/ by the three listener groups

Results of intermediate and advanced level learners found significant differences in the perception of /r/ under noise at SNR = 5, 0 dB (*t*-test, $p \leqq 0.05$). In the case of /l/, no significant differences were found between intermediate and advanced level learners' perception of /l/ in any of the listening conditions.

No significant differences were found between English native listeners and advanced level learners in the perception of /r/. The perception of /l/ by the two groups of listeners, however, found significant differences in all listening conditions (*t*-test, $p = 0.05$ in quiet, $p = 0.01$ in SNR = 10 dB, $p \leqq 0.01$ in SNR = 5 dB, and $p = 0.001$ in SNR = 0 dB).

Advanced level learners' performance had a dramatic decline from SNR = 5 dB (76.9%) to SNR = 0 dB (32.3%) in the perception of /l/. The decline at SNR = 0 dB was also observed in the native listeners – they scored 100% in all conditions in /l/ except under SNR = 0 dB at which their performance dropped to 73.3%.

Intermediate level learners' scores did not reach 70% even under quiet condition for the perception of /l/. The perception of /r/ by native listeners had no negative effect from the noise (all participants had perfect scores for all listening conditions). Advanced learners also had little effect from the background noise in the perception of /r/, regardless of its SNR. Intermediate level learners, however, were negatively affected by background noise, even under quiet condition.

4 Discussion

The present study reported the results of the perception of /r/ and /l/ in quiet and multi-speaker babble noise at SNR = 0, 5, 10 dB by Japanese native listeners with intermediate and advanced level English proficiency, and English native listeners. The correct rates showed that English native listeners were rarely affected by background noise for both /r/ and /l/.

The difficulty of the perception of /l/ compared to /r/ was observed in all three listener groups of the present study, a tendency that corresponds with native and non-native (Spanish) listeners of previous studies [2, 3]. Although native listeners' performance of /l/ in the present study reached 100% under quiet, SNR = 10 dB, and SNR = 5 dB, their performance dropped to 73.3% under SNR = 0 dB, whereas they scored 100% in all conditions in /r/. Advanced level learners' performance declined dramatically at SNR = 0 dB to around 30%. Intermediate level learners' performance also declined dramatically at SNR = 0 dB to around 30%, but their performance stays around 60% in the rest of the conditions including under quiet condition. This implies that advanced level learners need to be trained to perceive /l/ in SNR = 0 dB, while intermediate level learners need to be trained starting from quiet condition.

The results of the perception of /r/ in the present study, on the other hand, showed that advanced level learners had little difficulty with scores close to perfect in all listening conditions, which resembled the performance of native listeners. This implies that advanced level learners do not need much training of /r/ in either quiet or noise. Intermediate level learners' lowest scores for the perception of /r/ was at SNR = 5 dB (75.0%) and SNR = 0 dB (77.5%). With adequate training, they may improve their performance and become closer to near-native perception.

5 Conclusion

The present study showed that 1) advanced and intermediate level learners' performance differs, thus implying that perception training should be carried out at different levels of background noise according to the learners' proficiency, and 2) advanced level learners did not reach native-like performance especially in the perception of /l/ under background noise.

Akahane-Yamada *et al.* [8] showed in their experiment that perceptual training is effective for improving production, and the improvement obtained from the training is retained even after three to six months. Lecumberri & Cooke [13] also found that training is effective in consonant discrimination by non-native (Spanish) listeners, and that training in noise is more beneficial than training in quiet for voiceless consonants. The training effect of /r/ and /l/ in quiet and noise by Japanese listeners with different proficiencies is still unknown. If perceptual training is conducted under adequate conditions according to the listener's foreign language proficiency, namely under specific SNR conditions, the outcome should be more effective for each level of learners.

Acknowledgments

The authors would like to thank Professor Shigeaki Amano at Aichi Shukutoku University for his insightful comments. This study work was partially supported by Sophia University Open Research Center from MEXT.

References

[1] Takata & Nabelek, Acoust. Soc. Am., 88 (2), 663-666, 1990.

[2] Cutler *et al.*, J. Acoust. Soc. Am., 116 (6), 3668-3678, 2004.

[3] Garcia Lecumberri & Cooke, J. Acoust. Soc. Am., 119 (4): 2445-2454, 2006.

[4] Rogers *et al.*, Applied Psycholinguistics, 27, 465-485, 2006.

[5] Mayo *et al.*, JSLHR, 40, 686-693, 1997.

[6] Adachi *et al.*, Acoust. Sci. & Tech., 27, 5, 285-289, 2006.

[7] Ueda *et al.*, Acoust. Sci. & Tech., 28, 4, 251-259, 2007.

[8] Akahane-Yamada *et al.*, ISCLP, 606-609, 1996.

[9] TOEIC® (Test of English for International Communication), provided by ETS (Educational Testing Service).

[10] TOEFL® (Test of English as a Foreign Language), provided by ETS (Educational Testing Service).

[11] Varga & Steeneken, Speech Communication, 12 (3), 247-251, 1993.

[12] Boersma and Weenink, Praat: doing Phonetics by computer (Version 5.0.09) [Computer program].

[13] Lecumberri, M.L.G. & Cooke, M., JASA, 123 (5), 3881, 2008.