PERCEPTUAL TRAINING EFFECTS ON PRODUCTION OF ENGLISH /r/-/l/ BY JAPANESE SPEAKERS

Yasuaki Shinohara and Paul Iverson

University College London, UK y.shinohara@ucl.ac.uk; p.iverson@ucl.ac.uk

ABSTRACT

English /r/-/l/ perceptual training for Japanese speakers can improve production. However, it is unclear whether Japanese speakers actually improve their production of the primary acoustic cue for English /r/ and /l/, the third formant frequency (F3). The present study tested whether an English /r/-/l/ identification training program improves Japanese speakers' production of these phonemes in terms of F3. The results demonstrated that Japanese speakers lowered their F3 in English /r/ production after 5 training sessions, whereas F3 in /l/ production did not significantly change. This may be because F3 of Japanese lateral flap [r] is more similar to English /l/.

Keywords: second language, speech perception and production, training, improvement.

1. INTRODUCTION

It is well known that high-variability perceptual training for second language (L2) phoneme contrasts can improve the identification ability of its phonemes. Not only perception itself, but also production of L2 phonemes can be improved by perceptual training. Bradlow et al. [1] tested how English /r/-/l/ perceptual training for Japanese speakers affected their production, and they demonstrated that perceptual training improved both the perceptual ability of identification and production intelligibility. However, there has not been a study investigating what aspect of English /r/-/l/ production is improved by perceptual training. This study aimed to test whether perceptual training of English /r/-/l/ teaches Japanese speakers the distinction of the primary acoustic cue between English /r/ and /l/, the third formant frequency (F3), in terms of production.

2. METHOD

2.1. Subjects

A total of 20 native Japanese speakers (14 females; 6 males) completed 5 training sessions and pretest/posttest. Eleven speakers were tested in the UK and 9 speakers were tested in Japan. Their age ranged from 20 to 61 years old (median = 25.5 years old). They had no history of hearing impairment.

2.2. Training

Five training sessions of a computer-based English /r/-/l/ identification training program were given to subjects using their own laptops. Subjects listened to stimuli through headphones or earphones, and they were allowed to change the loudness to a comfortable level. They had one training session per day except one subject, and it was verified that all subjects completed all 5 sessions by checking all training logs. These logs were automatically recorded in their laptops and cannot be accessed by subjects due to protection by password.

The training task was perceptual identification of English /r/ and /l/. The stimuli were 50 wordinitial /r/-/l/ minimal pairs (100 words) repeated 3 times each, so that 300 trials were randomly given in each session. The stimuli were produced by 5 talkers and each session had a different talker. In the task, subjects saw a word-initial minimal pair on screen (e.g., *rock* vs. *lock*), heard a word, and clicked on a word which they thought they heard. When they had a correct answer, they heard a cash register sound and a repetition of the trial. If they had a wrong answer, they heard two beep sounds and two repetitions of the trial. It took approximately 30 minutes to complete a session.

2.3. Pretest/Posttest

Subjects took the same production tests before and after the 5 training sessions. The tasks were (1) reading 40 word-initial /r/-/l/ words (20 minimal-

pairs) in isolation and (2) reading the first third of The Rainbow Passage.

For word recordings, subjects read single words from a screen. The words and talkers used in pretest and posttest were not the same as used in the training stimuli. The word order was randomized, but it was same between pretest and posttest. In the passage reading task, subjects read the rainbow passage on the screen, and 7 wordinitial /r/ tokens and 6 word-initial /l/ tokens in the passage were analyzed.

2.4. Analysis

The F3 of /r/ and /l/ were measured for each token, and a linear mixed model with two fixed factors, testing block (pretest and posttest) and material (i.e., reading words in isolation and reading a passage) were used for the analysis. The random factors were subjects and pronounced words nested into subjects.

All 13 tokens from the reading passage task were used for the analysis, but for the task of reading words in isolation, 5 of 20 minimal-pairs were chosen for the analysis due to subjects' mispronunciation of the following vowels. In total, there were 920 tokens produced by 20 Japanese speakers, but F3 of 43 tokens were not able to be measured and were excluded from analysis.

3. RESULTS

Figure 1 displays the F3 values in English /r/ production including both tasks of reading words and reading a passage at pretest and posttest. The F3 values were normalised to the median F3 in the passage for the figure, but this step was not necessary for the statistical analysis. The linear mixed model demonstrated that there was a significant main effect of testing block, F(1, 228)= 20.68, p < .001, and the mean of F3 decreased by 107.81 Hz from an average of 2239.97 Hz at pretest to 2132.16 Hz at posttest. This result suggests that Japanese speakers learned the rhoticity of English /r/ after 5 training sessions.

Figure 2 shows the normalised F3 values in /r/ production for each material, reading words and a passage, at pretest and posttest. There was a significant main effect of material, F(1, 218) =46.17, p < .001. Across testing blocks, the F3 average of 2063.18 Hz in reading words in isolation was significantly lower than the F3 average of 2278.86 Hz in reading a passage, suggesting that subjects pronounced a more rhotic /r/ when they read English words in isolation. As displayed in Figure 2, the interaction between testing block and material was not significant, F(1, 228) = .18, p > .05, suggesting that subjects learned the rhoticity of English /r/ in both reading words and reading a passage.

Figure 3 displays the normalised F3 values in /l/ production of two different tasks, reading words in isolation and reading a passage. There was a significant main effect of material, F(1, 192) =6.59, p = .011, and the F3 average of 2709.77 Hz in reading words was significantly higher than the F3 average of 2595.92 Hz in reading a passage. This result suggests that subjects pronounced more distinctive /l/ when reading words than when reading a passage.

Figure 4 displays the normalised F3 values in English /l/ production of word reading and passage reading at pretest and posttest. There was no significant effect of testing block, F(1, 193) = 1.28, p > .05, or no significant interaction between testing block and material, F(1, 193) = .020, p > .05, suggesting that subject did not change F3 in /l/ production from pretest to posttest in either reading words in isolation or reading a passage.

4. **DISCUSSION**

Our results demonstrated that Japanese speakers improved their production of English /r/ in terms of F3 after 5 perceptual training sessions. The F3 in /l/ did not significantly change, but it is consistent to the prediction of Flege's Speech Learning Model (SLM; [2]). SLM predicts that it is more difficult to learn L2 phonemes when they are closer to an L1 phoneme category. Because Japanese lateral flap [r] is more similar to English /l/ than /r/ in terms of F3 [3], 5 training sessions may not be enough to change F3 of English /l/ production for Japanese speakers.

This result is in accord with SLM, but it is not consistent with the previous study [1]. Their result demonstrated that identifiability of both English /r/ and /l/ productions were significantly improved. The result of the improvement in identifiability of English /l/ production in the previous study may be due to Japanese speakers' improvement in the production of another acoustic cue such as transition duration and closure duration, but not F3.

Japanese speakers did not improve English /l/ production in terms of F3, but this may be explained by best exemplars of English /r/ and /l/ which Japanese speakers already had before training. Hattori and Iverson [3] demonstrated that Japanese speakers did not have native-like best exemplar of F3 for English /l/, although they have better perceptual exemplar of F3 for English /r/. This suggests that Japanese speakers might be able to change their production of F3 for English /r/ easily, because they already had its better exemplar in their mind. In case of English /l/, they may not have had native-like exemplar of F3 in English /l/, so that it may have been very difficult to improve the F3 for English /l/.

On the other hand, there seems to be another explanation for the smaller improvement in F3 for English /l/. There may not be enough space for the F3 improvement in the case of English /l/. For English /l/, the F3 produced by Japanese speakers is similar to English speakers' F3. However, F3 for English /r/ produced by Japanese speakers is not very similar to English speakers' production [3]. These differences suggest that the perceptual training would easily contribute to improve the production of English /r/, but it would be less effective for English /l/.

The current study's result also gives a clue to understand how Japanese speakers learn perceptual identification of English /r/ and /l/. Many studies been conducted to improve Japanese have speakers' perceptual identification ability of English /r/ and /l/, and Iverson et al. [4] concluded that Japanese speakers learn a strategy of labelling such that they label English /l/ if a stimulus is similar to their L1 lateral flap [r], and they identify English /r/ if it is not similar to Japanese flap [r]. However, if this conclusion is true, Japanese speakers cannot improve their production. In this study, it is demonstrated that subjects learned the rhoticity of English /r/. Therefore, it may be plausible to consider that Japanese speakers may have learned a lower level of acoustic perception as well as this strategy.

To investigate how Japanese speakers learn perception and production of English /r/ and /l/ through perceptual training, a future study testing the perception and production of more detailed acoustic cues is needed. Such a study could contribute to clarify the learning process of L2 phonemes and the link between perception and production.

Figure 1: Boxplots of normalised F3 in English /r/ production at pretest and posttest.

Normalised F3 (Hz) in /r/ Production



Figure 2: Boxplots of normalised F3 in /r/ production in reading words (white) and reading passage (grey) at pretest and posttest.



Normalised F3 (Hz) in /r/ Production

Figure 3: Boxplots of normalised F3 (Hz) in English /1/ production in reading words and in reading passage.



Normalised F3 (Hz) in /l/ Production

Figure 4: Boxplots of normalised F3 in English /l/ production in reading words (white) and reading passage (gray) at pretest and posttest.



Normalised F3 (Hz) in /I/ Production

5. REFERENCES

- Bradlow, A.R., Pisoni, D.B., Akahane-Yamada, R., Tohkura, Y. 1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual training on speech production. *J. Acoust. Soc. Am.* 101(4), 2299-2310.
- [2] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In: W. Strange. (eds), Speech Perception and Language Experience: Issues in Cross-Language Research. York: Baltimore, 233-277.

- [3] Hattori, K., Iverson, P. 2009. English /r/-/l/ category assimilation by Japanese adults: individual differences and the link to identification accuracy. J. Acoust. Soc. Am. 125(1), 469-479.
- [4] Iverson, P., Hazan, V., Bannister, K. 2005. Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *J. Acoust. Soc. Am.* 118(5), 3267-3278.