EFFECTS OF ENGLISH /r/-/l/ PERCEPTUAL TRAINING ON JAPANESE CHILDREN'S PRODUCTION

Yasuaki Shinohara & Paul Iverson

Faculty of Science and Engineering, Waseda University, Japan
Department of Speech, Hearing and Phonetic Sciences, University College London, UK
y.shinohara@aoni.waseda.jp; p.iverson@ucl.ac.uk

ABSTRACT

It has been proven that high-variability English /r/-/l/ perceptual training improves Japanese adults' production. The present study examined whether Japanese children also improve their production for the English /r/-/l/ contrast after perceptual training. The results demonstrated that production was significantly improved, but the improvement did not relate to an improvement in the primary acoustic cue use for the contrast (e.g., lowering F3 for English /r/ and raising F3 for English /l/). This suggests that children may respond to phonetic training differently from adults, and improve their production mostly through the modification of secondary acoustic cues.

Keywords: second-language, speech, perception, production, training.

1. INTRODUCTION

Second-language perceptual training has been successful for adults. For example, Japanese adults can improve their identification ability for the English /r/-/l/ contrast after high-variability identification training [1, 2, 3, 6, 7, 8], and perceptual training can improve production abilities [1, 2]. Children can also benefit from perceptual training; Japanese children aged 6-8 years improved their identification of English word-initial /r/ and /l/ through the course of 10 training sessions [10]. However, the perceptual training effects on Japanese children's production have yet to be examined.

The present study investigated whether Japanese children aged 6-12 years improve the identifiability of their English /r/-/l/ productions through the course of 10 perceptual training sessions. According to the motor theory of speech perception [4, 5], humans innately perceive speech sounds as articulatory gestures, suggesting that if children improve their perceptual identification as in Shinohara and Iverson [10], they may also improve their production, as Japanese adults did in previous studies [1, 2].

Another aim of the present study is to investigate how production abilities are improved after perceptual training. A previous study demonstrated that Japanese adults lowered F3 frequency for English /r/ but not English /l/ after five sessions of high-variability /r/-/l/ identification training [9]. In other words, Japanese adults improved the production distinction of English /r/-/l/ contrasts using the primary acoustic cue. The present study examined whether Japanese children also improve the English /r/-/l/ production distinction using this acoustic cue.

2. METHOD

2.1. Subjects

A total of 28 monolingual Japanese children aged 6-12 years (median: 8 years and 1 month) participated in this study. Of the 28 children, 16 (7 male and 9 female) were given 10 high-variability phonetic training sessions; 12 (6 male and 6 female) were not given any instruction between the pre and post tests and served as a control group. They had never lived outside of Japan and knew only some basic English such as greetings and color names. None had a history of language impairment.

2.2. Training

Ten computer-based perceptual training sessions were given to the children in the training group. Each training session consisted of three different perceptual tasks such as /r/-/l/ identification, auditory discrimination, and category discrimination. In the identification task (90 trials), English word-initial /r/-/l/ minimal-pair words (e.g., *rock-lock*) were displayed on a computer screen. Subjects heard minimal-pair words through headphones, and answered whether they heard /r/ or /l/. In the auditory discrimination task (48 trials), three identical objects were displayed on the screen, and subjects clicked on the one that had a different word-initial phoneme. The three stimuli were signal-processed English /r/-/l/ minimal-pair words manipulating only the primary acoustic cue (i.e., F3) difference between the /r/ and /l/ stimuli. This auditory discrimination task was intended to train Japanese children's F3 sensitivity at the /r/-/l/ phoneme boundary. In the category discrimination task (60 trials), three figures were displayed on screen, and subjects judged which one produced an English word starting with a different phoneme. The three stimuli for category discrimination were natural recordings of minimal-pair words produced by three different standard southern British English speakers.

Subjects received feedback in each training task. If they clicked on a correct answer, a message "せいかい" (*Correct*) and the highlighted answer were displayed on screen with a cash register sound. The stimulus was replayed for the identification task. When subjects clicked on a wrong answer, they saw a message "ざんねん" (*Bad Luck*) and a highlighted answer on the screen, and heard two descending beeps. The stimulus was replayed twice for the identification task and once for the auditory and category discrimination tasks. The stimuli used in training sessions were spoken by five standard southern British English speakers.

2.3. Pre/Post tests

Subjects completed pre and post tests before and after the training period. Japanese children pronounced 40 word-initial English /r/-/l/ minimal-pair words (20 pairs). Each English word was randomly displayed individually on a screen, and the subject pronounced them 2500ms after hearing an auditory prompt. During the 2500ms, they heard a sentence "Read the word after the tone" spoken by a male standard southern British English speaker and a beep sound.

For evaluating the productions, 10 native British English speakers heard the productions in a randomized order, and identified the phonemes, English /r/ or /l/. Since some Japanese children were not able to pronounce some English words, 28 tokens out of 2240 tokens were excluded from the analysis, leaving 2212 tokens. There was no feedback for the English subjects, but they were allowed to listen to the same trial as many times as they wanted.

F3 frequency was measured from the initial /r/-/l/ closure. Since Japanese children often pronounced apico-alveolar taps [ɾ] for English /r/ and /l/ and it was not possible to correctly measure F3 for these tokens, 10 English /r/-/l/ minimal pair words were selected for the analysis (e.g., *raw-law*, *reap-leap*, *rest-lest*, *root-loot*, *lug-rug*). Even so, 230 tokens out of the 560 tokens in total still needed to be excluded from the analysis, leaving 330 tokens.

3. RESULTS

Figure 1 displays the identifiability of English /r/ and /l/ produced by Japanese children at the pre and post tests. The production identifiability of the trained subjects improved from pre to post test, whereas there was no change in the identifiability of the untrained subjects in the control group.

The best fitting logistic mixed effects model was used for the production identifiability analysis. The model included the fixed factors of group (training vs. control), testing block (pre vs. post), and the interaction of group and testing block. Age was not included in the model because it did not have a significant effect on the model fit. The random factors were crossed intercepts of Japanese subjects and English word tokens. Neither the intercept of English subjects nor random slopes for testing block were included, because they did not contribute to the model's goodness of fit. The logistic mixed effect model demonstrated that although there was no significant main effect of testing block, $\chi^2(1) = 1.37$, $p > 0.05$, the main effect of group was significant, $\chi^2(1) = 8.96$, $p < 0.01$. There was also a significant interaction between group and testing block, $\chi^2(1) = 6.05$, $p = 0.014$, suggesting that Japanese children improved their production identifiability because of the high-variability training program.

Figure 2 displays the normalized F3 frequency of English /r/ and /l/ produced by Japanese children at pre and post tests. The best-fitting linear mixed effects model included the fixed factors of group, consonant (English /r/ vs. /l/), testing block and age. The five 2-way interactions and a 3-way interaction were also included in the model such as group and consonant, group and block, consonant and block, group and age, consonant and age, and group, consonant and block. The random factors were the same ones as included in the model for production identifiability analysis. The linear mixed effects model demonstrated that there was no significant effect of the fixed factors apart from the main effect of group, $\chi^2(1) = 4.00$, $p = 0.046$, and the interaction between group and age, $\chi^2(1) = 4.73$, $p = 0.030$. The 3-way interaction of group, consonant and block was also not significant, $\chi^2(1) = 0.70$, $p > 0.05$. This suggests that Japanese children did not improve their use of F3 after perceptual training.

Figure 3 displays the correlation between trained Japanese children's English /r/-/l/ production identifiability improvement rate and their F3 distinction change for the contrast from pre to post test. The F3 distinction change was calculated by subtracting the F3 change of English /r/ from that of English /l/ for each subject. After excluding four subjects from the analysis due to unmeasurable F3 values, a Pearson's correlation test was conducted. The Pearson's correlation test demonstrated that there was no significant correlation between the production identifiability improvement and F3 distinction change for the contrast, $r(10) = 0.41$, $p > 0.05$. This suggests that the production identifiability improvement did not significantly relate to the F3 distinction improvement.

Figure 1: Trained and untrained Japanese children's English /r/-/l/ production identifiability at pre and post tests, judged by native English speakers.

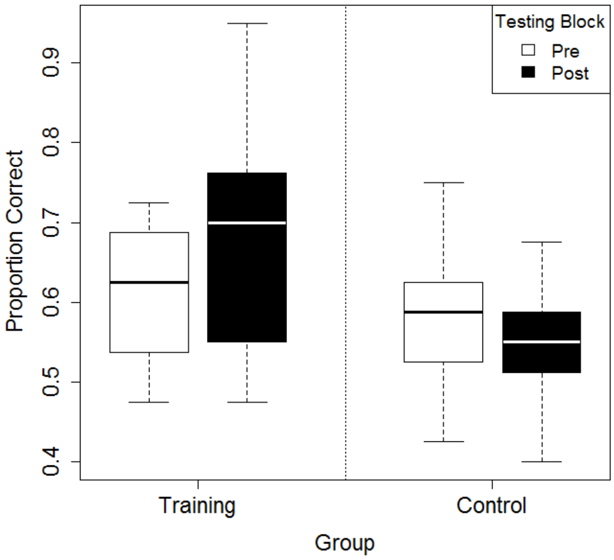


Figure 2: Trained and untrained Japanese children's F3 use for English /r/-/l/ production at pre and post tests.

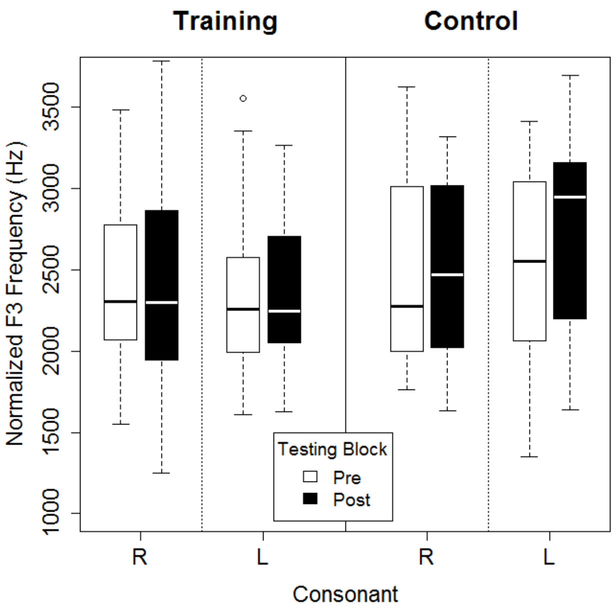
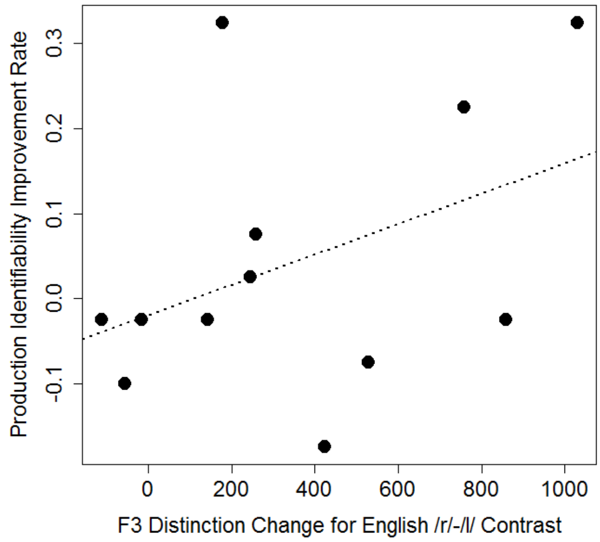


Figure 3: Scatterplot of the correlation between trained Japanese children's English /r/-/l/ production identifiability improvement rate and their F3 distinction change for the production contrast from pre to post test.



4. DISCUSSION

One of the main findings of the present study is that Japanese children significantly improve their production identifiability after intensive computer-based perceptual training. According to the motor theory of speech perception [4, 5], listeners innately perceive speech sounds as articulatory gestures. As hypothesized, if this perception-production link truly does not change through life span, children should also transfer their perceptual learning to production domains as do adults, and this was supported in the present study.

However, the production identifiability improvement seems not to relate to the improvement of the primary acoustic cue, F3. Since a previous study demonstrated that Japanese adults improve their primary acoustic cue use for the English /r/-/l/ contrast through perceptual training [9], it was hypothesized that children would also improve their F3 use for the contrast. Nevertheless, the present study demonstrated that Japanese children aged from 6-12 years did not significantly improve their F3 use for the English /r/-/l/ contrast.

Although the lack of F3 change seems to contradict the finding of production improvement, it might be that Japanese children improved their use of secondary acoustic cues, such as the duration of the transition or the closure. Such cues are generally not critical for native English speakers [3], but non-native aspects of articulations, such as rapid transitions, might make it harder for native English speakers to hear whatever F3 differences that the Japanese children make, particularly when the F3 distinctions are not clear.

There remains a question of why Japanese children did not improve the primary acoustic cue use in production. One plausible interpretation is that Japanese children had difficulties in pronouncing English /r/. English /r/ is often articulatorily characterised as bunched/curled tongue which often associates with lowering F3. Since native English children under 8 years old sometimes have difficulty in articulating the bunched/curled tongue for English /r/ [11], Japanese younger children in the present study may have likewise had difficulty.

Another interpretation is that Japanese children's perceptual learning is different from that of adults. Although this is inconsistent to the motor theory described above, it may be possible that some developmental factors such as phonemic awareness, cognitive attention and first-language acquisition may have contributed to the perceptual learning difference between adults and children. Consequently, their perceptual learning effects on production may have been different between adults and children.

5. ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number 26884062.

6. REFERENCES

[1] Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., Tohkura, Y. 1999. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & psychophysics* 61, 977–985.

[2] 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 learning on speech production. *J. Acoust. Soc. Am*. 101, 2299–2310.

[3] 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, 3267–3278. doi:10.1121/1.2062307

[4] Liberman, A. M., Mattingly, I. G. 1985. The motor theory of speech perception revised. *Cognition* 21, 1–36.

[5] Liberman, A. M., Cooper, F. S., Shankweiler, D. P., Studdert-Kennedy, M. 1967. Perception of the speech code. *Psychological Review* 74, 431–461.

[6] Lively, S. E., Logan, J. S., Pisoni, D. B. 1993. Training Japanese listeners to identify English /r/ and /l/. II: The role phonetic environment and talker variability in learning new perceptual categories. *J. Acoust. Soc. Am*. 94, 1242–1255.

[7] Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., Yamada, T. 1994. Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *J. Acoust. Soc. Am*. 96, 2076–2087.

[8] Logan, J. S., Lively, S. E., Pisoni, D. B. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *J. Acoust. Soc. Am*. 89, 874–886.

[9] Shinohara, Y., Iverson, P. 2013. Perceptual training effects on production of English /r/-/l/ by Japanese speakers. In J. Przedlacka, J. Maidment., & M. Ashby (Eds.), *Proceedings of the Phonetics Teaching and Learning Conference 2013* London, 83–86.

[10] Shinohara, Y., Iverson, P. 2013. Computer-based English /r/-/l/ perceptual training for Japanese children. *Proceedings of Meetings on Acoustics* 19, 060049 (2013).

[11] Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., Bird, A. 1990. The IOWA articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders* 55, 779–798.