

ACQUISITION OF THE ENGLISH /r/-/l/ CONTRAST BY JAPANESE SPEAKERS: EFFECTS OF TRAINING IN PERCEPTION ON PRODUCTION

Ann R. Bradlow and David B. Pisoni

Speech Research Laboratory, Indiana University, Bloomington, Indiana, U.S.A.

Reiko A. Yamada and Yoh'ichi Tohkura

ATR Human Information Processing Research Laboratories, Kyoto, Japan

ABSTRACT

This study investigates the effects of training in /r/-/l/ perceptual identification on /r/-/l/ production by adult Japanese speakers. Subjects were recorded producing English words that contrast /r/ and /l/ before and after participating in an extended period of /r/-/l/ identification training. Improvement in the Japanese trainees' /r/-/l/ productions as a consequence of training in perception was evaluated by a direct comparison of the pretest and post-test productions by a group of native American English listeners. The results showed significant perceptual learning for all subjects as a consequence of the training program. More importantly, this perceptual learning transferred to the production domain, implying a close link between perception and production.

INTRODUCTION

It is well known that the English /r/-/l/ contrast presents difficulties in both perception and production for Japanese speakers, for whom the contrast is neutralized [1]. Furthermore, this contrast has proved difficult to acquire in adult second language learners [2]. This finding has led to the claim that certain non-native phonetic contrasts may, in fact, be nearly impossible for adults to acquire [2]. However, this claim has recently been challenged by a new approach to perceptual learning [3]. Traditional methods of training non-native phonetic contrasts were guided by an attempt to draw the trainees' attention to the individual acoustic features that differentiate prototypical versions of the members of the target phonetic contrast. This approach has not been successful in promoting the acquisition of robust, English-like /r/ and /l/ categories by Japanese speakers [2]. In contrast, a new orientation that was designed to

expose the trainees to a wide range of exemplars of the target categories, has proven very successful [3]. The guiding principle behind this approach is that, in order to develop robust and linguistically meaningful phonetic categories, trainees must be exposed to exemplars that incorporate the variability that characterizes the target category. The success of this "high variability" training procedure suggests that adults are indeed capable of learning to perceive new, difficult phonetic contrasts [3].

The present study builds on this previous finding by investigating the effect of perceptual learning on /r/-/l/ production. Since the perceptual training program involves no production training whatsoever, transfer of the perceptual learning to the production domain would provide new evidence for a close link between perception and production, and would therefore be of both practical and theoretical interest.

METHOD

Perception Training

All perception training and testing was done at ATR Human Information Processing Research Laboratories. Eleven monolingual Japanese adults were trained over a period of 45 sessions using the "high-variability" training program. The stimuli consisted of English /r/-/l/ minimal pairs produced by five native English speakers. These minimal pairs included words with the target phoneme in multiple phonetic environments. Thus, the training stimuli incorporated a wide range of category variability due to cross-speaker differences, as well as differences in phonetic context. The procedure used for the training sessions was a two-alternative forced choice identification task, in which the trainees heard a stimulus and identified it from an /r/-/l/ minimal pair. (For example,

trainees heard "brush" and identified it from the pair, "brush-blush.").

In order to assess the trainees' improvement in identifying English /r/-/l/ minimal pairs, they performed a pre- and post-test at the start and end of the training period, respectively. In addition, the trainees performed two tests of generalization at the post-test phase of the experiment. These tests were designed to assess the extent to which the trainees could generalize the newly acquired /r/ and /l/ categories to stimuli they had not previously been exposed to. The first test of generalization presented new words produced by one of the talkers who produced the stimuli in the training set. The second test of generalization presented new words by a new talker.

Production Pre- and Post-test

In addition to the perception pre- and post-test, the trainees performed a production pre- and post-test. In this test, the trainees were recorded reading a list of English /r/-/l/ minimal pairs. The individual words were presented in random order and the subjects were given both visual prompts (standard English orthography) as well as an auditory model (a male, General American English speaker's production of the target word). The auditory model was provided in order to assist the Japanese trainees with the pronunciation of the rest of the word besides the /r/ or /l/. These trainee recordings were made in an anechoic chamber at ATR Human Information Processing Research Laboratories, and were digitized at a sampling rate of 22.05 KHz with 16 bit resolution.

Subjects

Eleven monolingual Japanese speakers (six males, and five females) served as subjects in the perceptual training program. A comparable group of eight Japanese speakers served as control subjects. None of the subjects had received any special English conversation training, although (as is typical in Japan) all had studied English since Junior High School (age 12 years). All subjects were recruited from Doshisha University, Kyoto prefecture, Japan.

The control subjects performed the perception pre- and post-tests, as well as the two tests of generalization; however, these subjects did not go through the

training program. In addition to the perception tests, these control subjects also performed the production pre- and post-tests. The time lag between the control subjects' pre- and post-tests was identical to the time of the training program for the experimental subjects.

Evaluation of Trainee Productions

The Japanese trainees' pre- and post-test utterances were transferred to the Speech Research Laboratory at Indiana University, where they were converted to 12 bit resolution for presentation to American English (AE) listeners using a PDP-11 computer. The AE listeners performed a direct comparison task on pairs of pre- and post-test tokens. In this task, the AE listeners saw the target word in standard English orthography on a CRT monitor and then heard two versions (a pre-test and a post-test version) of the target word by a single Japanese trainee. The AE listeners responded by selecting the version that sounded "better," or "more precisely articulated" on a seven-point rating scale. On this scale a response of "1" indicated that the first version was "much better" than the second version; a "7" indicated that the second version was "much better" than the first version, and a "4" indicated that there was no difference between the two versions. In the presentation of these stimuli, each pretest - post-test pair was presented twice: once in each of the two possible orders (pretest then post-test, and vice versa).

Each Japanese trainee's pre- and post-test productions were compared by a separate group of ten AE listeners, for a total of 110 AE listeners. Control subjects' pre- and post-test productions were evaluated by an additional 80 AE listeners (ten for each of the eight control subjects). These AE subjects were all students at Indiana University and received course credit for their participation in this experiment. All subjects reported no history of speech or hearing impairment.

RESULTS

Perceptual Learning

Figure 1 shows the Japanese trainees percent correct identification for the pretest, post-test, and the two tests of generalization. The left panel shows the

results for the trained group, and the right panel shows the results for the control group. For the trained group, there was a significant improvement in accuracy from pretest to post-test ($t(10)=-7.38$, $p<.001$ by a 2-tail paired t-test), and this level of performance was maintained in the two tests of generalization. In contrast, the control group showed no difference from pretest to post-test ($t(7)=2.185$, $p=.065$ by a 2-tail paired t-test).

These data replicate the results of previous /r-/l/ perception training studies using the "high-variability" training procedure [3]. This pattern of results indicates that the trainees did indeed show significant perceptual learning as reflected in the significant changes in performance for the experimental group. In the present study, we were also interested in investigating how this perceptual learning affected the subjects' ability to produce more native-sounding words that contrast in /r/ and /l/.

Transfer of Perceptual Learning to Production

Figure 2 shows the distribution of responses for the AE listeners' comparisons of the Japanese trainees' (left panel) and control subjects' (right panel) pre- and post-test productions. This figure shows the proportion of trials for which the AE listeners judged there to be no difference between the pre- and post-test productions (post=pre), for which they judged the pretest version better than the post-test version (post<pre), and vice versa (post>pre).

For both the trained and control subjects, there was a relatively small proportion of trials that received the post=pre response. This proportion was higher for the control subjects than for the trained subjects. More importantly, a far greater percentage of the trained subjects' productions received a post>pre rating than the reverse rating. This is seen in Figure 2 by the significant difference between the frequency of post<pre and post>pre responses for the trained subjects' productions ($t(10)=-3.018$, $p=.013$ by a 2-tail paired t-test). In contrast, for the control subjects, there is no difference in frequency of the post<pre and post>pre responses ($t(7)=-.625$, $p=.552$ by a 2-tail paired t-test). In

other words, the AE listeners showed a preference for the trained subjects' post-test productions over their pretest productions; whereas, the AE listeners showed no such preference for the control subjects' post-test productions.

Thus, the post-test /r-/l/ utterances of the Japanese subjects who went through the training program showed significant improvement over the corresponding pretest productions. This result demonstrates a transfer to the production domain of the perceptual knowledge that was acquired during the training program.

DISCUSSION

The results of this study show that, even with no explicit production training, the perceptual learning that resulted from the /r-/l/ identification training transferred to the production of /r/ and /l/. From a practical point of view, this finding suggests that the acquisition of new phonetic contrasts in production, as well as in perception, can be facilitated by extensive training in perception alone. From a theoretical point of view, these data indicate a close perception-production link, to the extent that learning in the one domain transfers to changes in the other domain.

ACKNOWLEDGMENTS

We are grateful to Takahiro Adachi for technical support, and to Fernando Vanegas for subject running. This work was supported by NIDCD Training Grant DC-00012 and by NIDCD Research Grant DC-00111 to Indiana University.

REFERENCES

- [1] Goto, H. (1971), "Auditory perception by normal Japanese adults of the sounds 'l' and 'r,'" *Neuropsychologia* vol.9, pp. 317-323.
- [2] Strange, W., & Dittman, S. (1984), "Effects of discrimination training on the perception of /r-/l/ by Japanese adults learning English," *Perception and Psychophysics* vol. 36, pp. 131-145.
- [3] Lively, S., Pisoni, D., Yamada, R., 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.* vol. 89, pp. 874-886.

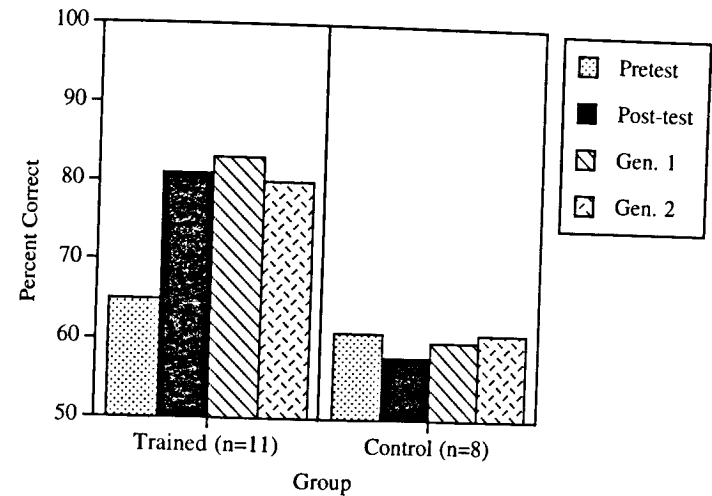


Figure 1. Japanese trained and control subjects' performance on the perceptual identification pretest, post-test and two tests of generalization.

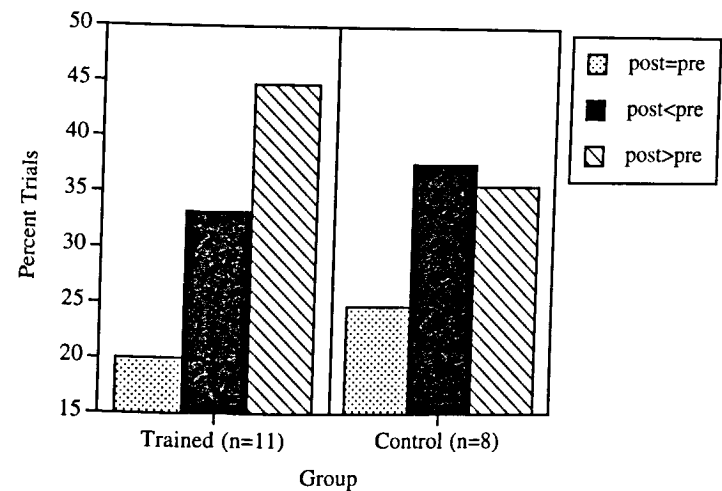


Figure 2. Distribution of AE listener comparison responses for the pre- and post-test productions by trained and control Japanese subjects.