Modification of L2 Vowel Production by Perception Training as Evaluated by Acoustic Analysis and Native Speakers

Reiko Akahane-Yamada*  Winifred Strange†  Jesica C. Downs-Pruitt*
Yasushi Masuda*‡

*ATR Human Information Processing Research Laboratories, Kyoto, 619-0288, Japan
†University of South Florida, Tampa FL 33620-8150
‡Graduate School of Information Science, NAIST, Nara, 630-01, Japan

Abstract: To examine the relationship between speech perception and production in second language (L2) acquisition, this study investigated the effects of L2 vowel perception training on L2 vowel production. Native speakers of Japanese served as subjects and received a pretest, identification training on English /a, æ, ə/ and a post-test. A recording of the production of these vowels was also made at pre- and post-test. Identification ability significantly improved from pretest to post-test. However, the effect did not transfer to the production domain.

1 INTRODUCTION

Recent cross-language studies have shown that there is a link between speech perception and speech production when learning L2 consonants. As clear evidence, Bradlow et al. [1] demonstrated that identification training of the English contrast /r/ and /l/ improved Japanese speakers' production ability significantly. Recently, Strange and Akahane-Yamada [2] found that identification training of English vowels improved Japanese speakers' identification ability dramatically (from 42% in pretest to 85% in post-test). The current study examined whether this vowel identification training also transferred to production ability.

2 EXPERIMENT

2.1 Subjects

Twenty-two native speakers of Japanese served as subjects. Twelve subjects were randomly selected and assigned to the trained group, which received identification training. The remaining 10 subjects were assigned to the control group and participated only in the pre- and post-tests.

2.2 Procedure

The experiment employed a pretest/post-test design. The trained group received 36 sessions of identification training on /a, æ, ə/ over 9 days. Before and after the training period, recordings were made of the subjects' productions of /æ, e, ə, / in several CVC syllables in both citation and sentence form. The productions were later digitized and evaluated in two ways, first by acoustical analysis and second by native speakers of American English (AE). As a comparison, a recording of the same speech materials was made of 11 native speakers of AE.

In the native speakers' evaluation, a panel of nine native speakers of AE judged pretest and post-test versions of vowels produced by subjects in the training and control group. In the first session, the intelligibility of the vowels was assessed. The sequence of the stimulus presentation was blocked by subject; the pre- and post-test versions from one subject was presented in a single block in random order. The evaluator identified the vowels by selecting one out of a possible 13 AE vowels. Half of the listeners evaluated citation form first followed by sentence form, and the other half of the listeners evaluated sentence form first followed by citation form. In the second evaluation session, evaluators rated the goodness of the productions with knowledge of the talkers' intended vowel using a 7 point scale (1: worst - 7: best).
The productions from the three groups (training, control, native AE) were acoustically analyzed; formant frequencies at the mid-point of the vowel period were extracted.

![Figure 1](image)

**FIGURE 1:** F1 and F2 frequencies for the five AE vowels, eh(r), ah(ʌ), ao(ɔ), aa(ʌ), and ae(ɛ). Left panel shows the AE speakers' productions by speech style (citation and sentence). The other two panels show the Japanese speakers' productions of the same vowels by phase of the recording (pretest and post-test) in citation form (middle) and sentence form (right).

### 3 RESULTS

From the human evaluation, intelligibility scores (in terms of how often the AE listeners' judgments matched the talkers' intended vowels) did not improve significantly from pretest to post-test in either the trained group (37% to 40%) or the control group (33% to 36%). The rating score changed significantly from pretest to post-test but in both groups (4.2 to 4.5 vs. 4.4 to 4.6).

In the acoustical analysis, the vowel distributions on F1-F2 plane were compared to the native AE speakers' productions (Figure 1). A large difference between Japanese speakers and AE speakers was observed. We hypothesized that Japanese speakers would assimilate English vowels into their native language categories in production, as well as in perception (reported by Strange et al., 1997). However, again, we found no significant difference from pretest to post-test.

### 4 DISCUSSION

Compared to the /r-l/ training, there was a larger improvement in perception ability. However, vowel production ability did not improve substantially as it did for /r-l/. This finding suggests that the vowel distinction in production is more difficult to acquire than for consonants because their articulatory control is continuous while the articulatory control of consonants is somewhat discrete.

### 5 REFERENCES