Influence of aging on perceptual learning of English phonetic contrasts by native speakers of Japanese

Rieko Kubo* and Reiko Akahane-Yamada

ATR Human Information Science Laboratories,
2–2–2 Hikaridai, “Keihanna Science City,” Kyoto, 619–0288 Japan

(Received 7 September 2005, Accepted for publication 22 September 2005)

Keywords: Second language, Perception training, Speech perception, Aging

PACS number: 43.71.Lz, 43.71.Hw [DOI: 10.1250/ast.27.59]

1. Introduction

The effect of age in acquiring phonetic categories in a second language (L2) has been examined by two approaches. First approach is to examine the effect of the onset age of exposure to an L2-speaking environment on the acquisition of perception and/or production of new phonemes that do not occur in the subjects’ native language. For example, Flege et al. [1] examined immigrants’ acquisition of new phonetic categories, and Akahane-Yamada [2] examined native speakers of Japanese who had once been exposed to an English-speaking environment and analyzed their acquisition of the English sounds /r/ and /l/. Recently, another approach has been introduced. In this second approach, laboratory training is conducted, and the relationship between age of the trainee and the improvement through training is analyzed (e.g. [3]). However, no study has yet examined the training effect in elderly people. In this paper, we have trained two groups of native speakers of Japanese - young adults and elderly adults, and compared the training effects between the two groups. Questions addressed are: 1) does the training improve elderly adults’ ability to identify L2; and if it does, 2) is there any difference in the extent of improvement between elderly and young adults.

2. Experiment

Stimuli English words minimally contrasting in /r/-/l/ (RL words), /b/-/v/ (BV words) and /s/-/z/ (STH words) were employed as word materials. Twenty pairs of RL words, ten pairs of BV words and ten pairs of STH words were used in the training. Each of these words was produced by five native speakers of American English (three males and two females). Another twenty pairs of RL words, ten pairs of BV words and ten pairs of STH words were used in the pretest, mid-tests and post-test. Each word was produced by another native speaker of American English (male). Recordings were made in an anechoic chamber at ATR laboratories, and individual words were saved into separate audio files at 44.1 kHz sampling frequency and 16-bit resolution.

Participants Two groups of native speakers of Japanese were tested: 19 young adults ranging in age from 18 to 26 years old, 20.3 on average (group YA), and 22 middle-aged and elderly adults ranging in age from 53 to 70 years old, 59.2 on average (group ME). None of them had reported having lived abroad for more than three months. In addition, none of them reported any history of hearing or speech disorders. A hearing test was performed for the frequencies from 250 to 8,000 Hz using an audiometer (AA-61BN). The signal softer than 15 dB Hearing Level was not measured. Results averaged over participants within each age group are shown in Fig. 1. It is clear that the elderly adults showed significant hearing loss at higher frequencies.

Procedure A pretest/post-test design used by Strange and Dittmann [4] was modified and employed as the experimental design. In Strange and Dittmann, the pretest and post-test were administered before and after the training period, but in the current study, mid-tests were also administered during the training period. In all the tests, a two-alternative forced-choice (2AFC) task was used. During each trial, a member of a minimal pair was played binaurally through the headphones (ATH-AD5) at a comfortable listening level (which was maintained throughout the experiment). Participants were asked to identify one of the members of the minimal pair by selecting one of the orthographic representations of them on a CRT monitor. There was no feedback to the subjects on their responses.

Each test consisted of three blocks. In the first block, RL words were presented in a random order for a total of 40 trials. BV words and STH words were presented in the second and the third block, respectively, in random order, giving 20 trials in each block.

The same task used in the tests was used in the training blocks. However, in this case feedback was provided to the participants on their responses. A chime sounded for a correct answer and a buzzer sounded for an incorrect answer. Forty RL words, 20 BV words and 20 STH words produced by one talker were presented in one training session. One training unit consisted of five training sessions, and five talkers cycled in a fixed order.

A pretest was administered at the beginning of the experiment, in which tests were administered for every three training units (i.e., 15 training sessions). Twenty-seven training units were completed, resulting in ten tests. The tests were administered individually at ATR laboratories. Participants in group ME brought a laptop PC and headphones to their home and conducted training sessions in a quiet room. Participants in group YA received training sessions at ATR laboratories.

* e-mail: rkubo@atr.jp
3. Results

During the Training  Correct response rate for each contrast in each block for each participant was calculated, and then they were averaged over the participants in each age group. Figure 2 shows that both groups improved during training.

During the Test  The correct response rates for each phonetic contrast, test and age group were calculated (Fig. 3). The arc-sine-transformed values were submitted to a three-factor analysis of variance (ANOVA), where the test phase (test1–test10) and phonetic contrasts (RL, BV, STH) were within-subjects variables, and age group (YA and ME) was a between-subject variable. The main effect of phonetic contrast \( F(2,78) = 144.78, p < 0.001 \) and test phase \( F(9,351) = 48.12, p < 0.001 \), and interactions between age group and phonetic contrast \( F(2,78) = 16.62, p < 0.001 \), and between phonetic contrast and test phase \( F(18,702) = 4.935, p < 0.001 \) were significant. Other effects were not significant. Results of a multiple comparison indicated that test1 and test10 differed significantly, and the accuracy of group ME was significantly lower than that of group YA only for contrast BV, but not for contrasts RL or STH.
4. Discussion
The present results demonstrated that perceptual training improves the ability to identify difficult phonetic contrasts in L2, even for middle-aged and elderly adults (50’s to 70’s), and the extent of the improvement does not substantially differ between young adults (20’s) and the middle-aged and elderly (50’s to 70’s).

It was expected that hearing loss at higher frequencies would prevent trainees from improving their ability to perceive /s/-/θ/ and /b/-/v/ contrasts, because the dominant acoustic characteristics of /b/, /v/, /s/, and /θ/ are non-periodic noises with rather broad spectra: a burst for /b/ and a fricative for /v/, /s/-/θ/. However, the result indicated that middle-aged and elderly participants improved their ability to perceive these contrasts even though they had hearing loss at higher frequencies. The results suggest that the middle-aged and elderly adults focused on other cues to perceptually differentiate those contrasts. Similar phenomena have been reported whereby hearing-impaired people identify /s/-/θ/ sounds [5] and that native English speakers’ perception of /b/-/v/ and /s/-/θ/ was hardly distorted by noise [6].

The present paper is an interim report on a study examining the effect of age when native speakers of Japanese learn to perceive English /t/-/l/, /b/-/v/ and /s/-/θ/ sounds. Extended training has also been given to trainees spanning a wider range of ages, from 10 years old to 70 years old. The result of that study will further clarify the effect of age in acquiring speech perception of a second language.

Acknowledgments
This research was supported in part by the National Institute of Information and Communication Technology of Japan.

References