Training American Listeners to Perceive Mandarin Tones: A First Report

Yue Wang, Allard Jongman, and Joan A. Sereno

Cornell Phonetics Laboratory, Cornell University, Ithaca, New York 14853

Abstract: Auditory training has been shown to be effective in the identification of nonnative segmental distinctions. This study attempted to investigate whether such training is applicable to the acquisition of nonnative suprasegmental contrasts, i.e., Mandarin tones. In four sessions lasting a week, four American learners of Mandarin were trained to identify the four tones in natural words. The trainees' identification accuracy showed an overall 11% increase from the pre- to the posttest, indicating that tone contrasts can be improved by training. The results are discussed in terms of nonnative perceptual modification at the suprasegmental level, and some methodological implications for further studies.

Previous research has shown substantial improvements in the identification of segmental distinctions which are absent in the listener's native language after simple phonetic laboratory training procedures (e.g., 1, 2, 3, 4). However, little research has reported the application of such training procedures to the acquisition of nonnative speech contrasts at the suprasegmental level. In the present study, American listeners were trained to perceive Mandarin lexical tones, which have been observed to be difficult to acquire for learners whose native language is nontonal (e.g., 5). Mandarin has four distinctive tones, with Tone 1 having high-level pitch, Tone 2 high-rising pitch, Tone 3 low-dipping pitch, and Tone 4 high-falling pitch. The goal of this study was to examine whether auditory training which has been shown to be effective at the segmental level is applicable to the acquisition of nonnative suprasegmental contrasts.

Eight native listeners of American English participated in the study voluntarily, with four as trainees, and four as control subjects. All had been learning Mandarin for one and a half years at the time of the study. Three native speakers of Mandarin Chinese participated as talkers, with one female reading the pre/posttest stimuli, and two others (one male and one female) providing the training stimuli. The stimuli were real isolated Mandarin words, including 100 items (25 for each tone) for the pre/posttest, and 160 (40 for each tone) for training. The training program adopted the high-variability procedure shown in previous studies to be efficient to promote the formation of robust nonnative phonetic categories (1). This procedure consisted of a pretest, a training phase, and a posttest. Both the trained and the control subjects took the pretest, in which they were presented with the 100 randomized stimuli (in pinyin) on a sheet of paper, and were to mark, for each item, which of the four tones they had heard. In four sessions lasting a week (40 minutes per session), the trainees were presented with tape-recorded stimuli produced by either of two talkers. The four tones were trained pairwise (i.e., Tones 1 and 2, Tones 1 and 3, Tones 1 and 4, Tones 2 and 3, Tones 2 and 4, and Tones 3 and 4). The trainees' task was two-alternative forced choice identification, with immediate feedback after each trial. Both the trained and the control subjects took the posttest, which was otherwise identical to the pretest, except that the stimuli were rerandomized.

Percent correct identifications in the pre- and posttest were compared to determine if tone contrasts were improved by training. The overall results (Figure 1) reveal an 11% increase in identification accuracy for the trained group, as opposed to only 2% for the control group. A comparison of the trainees' identification accuracy for pretest and posttest shows an increase for all subjects (Figure 2) and for all tones (Figure 3). An analysis of tone confusions (Figure 4) indicates that the most easily confused tone pair was Tones 2 and 3, followed by Tones 1 and 2. In sum, Figures 1 to 4 indicate unanimously that the trainees' identification of the four tones improved after training.

The results of the present experiment suggest that the perception of Mandarin tones can be improved using a simple training task. This indicates that the procedure which has been adopted in training the acquisition of nonnative segmental contrasts can also be applied at the suprasegmental level. These preliminary results have several implications for future studies. First, the results showed an overall 11% increase in the trainees' identification accuracy after training, which is comparable to the results of other training studies (1, 2, 3, 4). This is indeed a substantial improvement considering the short period of training. Given that previous studies have found progressive improvement in the trainees' performance as training continued (1), a greater degree of improvement might be expected if tone training lasted longer. Second, the results suggest that the high-variability paradigm is also effective in tone training. In fact, talker variability is crucial in tone training, since different talkers (especially males and females) have different fundamental frequencies (F0). Since native Mandarin speakers use changes in F0 contours more than height to distinguish among tones (e.g., 6), if various talkers are used (including both males and females), learners will be trained to focus on detecting the pitch contour differences of the tones, and to normalize the
differences in F0 height of various talkers. Third, the results also showed that the most easily confused tone pair was Tones 2 and 3, followed by Tones 1 and 2, and Tones 1 and 4. This strongly suggests that for future studies, training should be dedicated more to those difficult tone pairs. Finally, this study only examined the effect of tone training by comparing the performance at pre- and posttest. A future study will examine generalization and long-term retention of the training stimuli, given that previous studies have shown that segmental contrasts obtained in training can be generalized to novel words and talkers (2), can be maintained long after training (3), and learning gained perceptually can be transferred to production (4).

In sum, the present results suggest that auditory training can be used to modify American listeners' perception of Mandarin tones in isolated words. Based on the present results, a large-scale study is currently underway with more subjects and talkers to examine the training effects on the acquisition of nonnative suprasegmental contrasts, and the transfer, generalization, and long-term retention of the contrasts gained from training.

REFERENCES