2aSC13. Effects of phonological training on tone perception for English listeners

Chang Liu* and Bharath Chandrasekaran

*Corresponding author's address: Communication Sciences and Disorders, The University of Texas at Austin, CMA A4.130, Austin, Texas 78712, changliu@utexas.edu

The goal of this study was to examine the extent to which phonological training improves categorical perception of Mandarin Chinese tones in native speakers of American English. Two sets of F0 continuums were generated from rising to level tones and from falling to level tones. Participants underwent identification tasks before (pre-training) and after (post-training) phonological training. Pre-training tests showed that the tone identification shifted gradually from contoured tones (rising/falling) to level tones as a function of F0 frequency. Phonological trainings were provided to listeners in two consecutive days with each session lasting for one hour. In phonological training, listeners were provided immediate feedback (correct/incorrect) after making response in Mandarin tone patterns. Results showed that tone identification function became significantly steeper with phonological training. Ongoing work is now testing the extent to which a longer training regiment can enhance categorical perception of tones.

Published by the Acoustical Society of America through the American Institute of Physics
1. INTRODUCTION

It is well established that native English speakers have difficulty in categorizing lexical tones, a factor that could contribute to difficulties in learning tone languages (Wang et al. 1999). Perception of lexical tones for English speakers is critically dependent on their psychophysical capacity to resolve pitch movement (Xu et al., 2006; Wu and Lin, 2008; Peng et al., 2010; Liu, 2013). In particular, for tone identification, English listeners have shallower slope and different tonal boundary, compared to listeners with a tone language background (Xu et al., 2006). The extent to which categorical boundaries can be modulated by short-term training has not been systematically explored for tonal contrast. In the present study, we examine whether and how phonological training of tone identification improves English listeners’ formation of lexical tone categories.

Wang et al. (1999) trained American English listeners to perceive lexical tones. English listeners received auditory training of four naturally-produced lexical tones with eight sessions in two weeks. They found that Mandarin tone identification scores increased by more than 20% for the post-training test compared to the scores of the pre-training test. In the present study, we measured tone identification using a categorical perception approach (i.e., tone categorization for a series of stimuli with F0 continuum) before and after auditory training of lexical tones. The goal of this experiment was to examine the extent to which phonological training improves English listeners’ categorical perception of lexical tones.

2. METHODS

2.1. Stimuli

For the categorical perception experiment, two sets of target stimuli were constructed based on a Mandarin Tone 1 (phonetically described as ‘high-level’) word [ma1] “mother”, recorded from a male native speaker of Mandarin (the first author) with normal speech and hearing function. The steady portion of the recorded [ma] with a duration of 200 ms was selected as the standard stimulus for F0 manipulation. The onset and offset of F0 frequency were modified to be 150 Hz. One set of stimuli was a rising-to-level continuum, constructed by manipulating the onset F0 of the recorded word “mother” from 100 to 150 Hz. The offset F0 were manipulated from 150 to 100 Hz to form the other falling-to-level continuum. Both continua have 11 equally spaced steps at 5 Hz. The speech synthesizes were manipulated in STRAIGHT (Kawahara et al., 1999), a speech synthesizer using a pitch-adaptive method to generate high-fidelity speech (Liu and Kewley-Port, 2004; Assmann and Katz, 2005). For the phonological training, stimuli consisted of native production of the four Mandarin tones: Tone 1 (T1), tone 2 (T2), tone 3 (T3), and tone 4 (T4). Mandarin Chinese words (bu, di, lu, ma, and mi) that are minimally contrasted by tone categories were used in the category training experiments. By using different segments and multiple talkers, our aim was to expose learners to high variability present in natural language. Each of these syllables was produced in citation form with the four Mandarin tones. Talkers consisted of 4 male and 4 female native speakers of Mandarin Chinese originally from Beijing. Each talker produced four tokens...
of each stimulus. Stimuli were RMS amplitude and duration normalized (70 dB SPL, 0.44 s) using the software Praat.

2.2. Participants

Six young English Monolingual speakers (20-22 years old) with normal hearing participated in this study. All participants had no learning experience of tone languages. No participant received formal five-year musical education and none had regular musical performance in recent three years.

2.3. Procedures

Each participant was tested in sound-treated booths. Stimuli were presented by a TDT digital sound processor (RP 2.1) and a headphone buffer (HB 7.1) to the right ears of the participants at 70 dB SPL through ER-2 insert earphones. The sound-pressure level measured in the NBS-9A 2-c^3 coupler by a Larson-Davis sound-level meter (Model 2800) with the linear weighting band set at 70 dB SPL for the speech signal. A ten-minute training session was used before the tests to familiarize participants with the experimental procedure for the identification task. Short breaks were provided as needed.

For the identification experiment, listeners were asked to categorize the stimulus as a falling, flat, or, rising tone. The target stimuli from each of the continua were presented in isolation with 20 repetitions in a random order. The rate of presentation was self-paced. The next stimulus was presented automatically in 1 second after a response was collected.

2.4. Training sessions

There were two training sessions in consecutive days. Each phonological training session followed the test session (e.g., tone identification) on daily experiments. During the phonological training session, participants were asked to categorize stimuli into one of four categories. Feedback was provided to reduce tone errors. After each stimulus presentation, listeners’ task was to identify the lexical tone as rising, flat, falling-then-rising, or falling tones. Following listener’s response, a feedback was provided to indicate the correctness of the response.

2.5. Data analysis

Identification scores of Tone 1 in percentage were measured at each onset frequency (for Tone 1-2) and offset frequency (for Tone 1-4). For example, for Tone 1-2 identification, when onset frequency of the stimuli is 140 Hz, if the participant considered 16 out of the 20 repetitions of the stimuli as Tone 1, the identification score for this participant at 140 Hz was 80%. Based on the binomial distribution of the identification scores and the sigmoid shape of the response function, a sigmoid model Eq. (1) between the identification score (y) and the onset/offset frequency (x) was fitted to obtain the identification function for each listener. Goodness of fit was also evaluated to determine whether identification performance in each listener fit the sigmoid model.

\[
y = \frac{1}{1 + e^{-\frac{x-x_0}{b}}}
\]

(1)
For each fitting function, b and x₀ are the regression coefficients estimated by sigmoid model, representing sharpness and categorical boundary of the identification function. The fitness of the sigmoid model for tone identification was conducted in SigmaPlot® v10.0 that used the Marquardt-Levenberg algorithm (Marquardt, 1963) to find the parameters of the independent variable (x), giving the best fit between the equation and the data. An iterative process was conducted for the algorithm to seek the values of the parameters that minimized the sum of the squared differences between the values of the observed and predicted values of the dependent variable (y).

To investigate the effects of language experience on identification performance, we obtained individual indices for participants whose identification performance yielded significant fitting for the sigmoid model. For each listener in each tone continuum type, categorical boundary x₀ was calculated from the value of onset/offset F₀ when identification score y equals to 50%, while the sharpness of categorical boundary was represented by the parameter b in Eq. 1.
3. RESULTS

Figure 1 illustrates average identification scores of tone 1 over the six listeners as a function of the onset F0 frequency for three consecutive days, while Figure 2 shows the corresponding fitting sigmoidal functions. A two-factor (F0 continuum x training day) repeated-measures analysis of variances (ANOVA) was conducted with x0 (boundary) and b (sharpness) as the dependent variable, respectively. Results showed that for both x0 and b, training day had a significant effect (x0: $F_{2,10} = 4.619$, $p < 0.05$; b: $F_{2,10} = 5.570$, $p < 0.05$), while F0 continuum (rising and falling) had no significant effect as well as the two-factor interaction (all $p > 0.05$). Tukey post hoc tests indicated that the identification function became significantly steeper for post-training sessions ($p < 0.05$), i.e., greater values of b for day 2 and 3 (post-training) than for day 1 (pre-training).
4. DISCUSSION

As shown in Fig. 1, before phonological training, native English listeners had difficulty to categorize the three lexical tones, e.g., shallow shift of tonal categorization from one category to another as F0 frequency changed. However, after phonological training, native English listeners were able to perceive tones of speech stimuli more categorically, i.e., steeper slopes in the identification function. These results suggest that training with naturally produced Chinese speech stimuli may help English listeners shift their categorical representations. Our results are consistent with the finding that auditory training enhances English listeners’ accuracy of lexical tone identification (Wang et al., 1999). Future research is to focus on developing training methods (lexical vs. phonological training) and explore the functional ramifications of individual differences in categorical perception (Xu et al., 2006; Peng et al., 2010).

REFERENCES