Effects of tone training on Cantonese tone-word learning

Angela Cooper and Yue Wang
Department of Linguistics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada
akcooper@u.northwestern.edu, yuew@sfu.ca

Abstract: The present study examined the effect of improving lexical tone identification abilities on Cantonese tone-word learning. Native English non-musicians received training on Cantonese tones before learning the meanings of words distinguished by these tones. Their results were compared to English non-musicians and musicians who received no tone training. The tone-trainees obtained a similar level of word identification proficiency as musicians by the end of training and were significantly better than non-tone trained non-musicians. These results lend support for phonetic-phonological-lexical continuity in learning because enhancing listeners’ perception of lower-level tonal information significantly contributed to success in a higher-level linguistic task.

1. Introduction
Long-term pitch experience, both linguistic and non-linguistic (e.g., musical training), has been shown to significantly affect the perception of lexical tone contrasts as well as the ability to utilize such contrasts to make lexical distinctions. Previous research has suggested that long-term exposure to native tonal contrasts or non-linguistic musical pitch aids listeners in better attuning to the pitch modulations of non-native tonal distinctions, as both tone-language listeners and musicians identified and discriminated lexical tones more accurately than non-tone-language non-musicians (Lee and Hung, 2008; Wayland and Guion, 2004). Moreover, short-term pitch experience (e.g., tone training) can also facilitate the perception of non-native tonal contrasts (Francis et al., 2008; Wang et al., 1999). For example, Wang et al. (1999, 2003) reported that after training American English listeners to identify Mandarin tones, they not only improved their tone identification accuracy but also generalized these improvements to new talkers, stimuli, and the production domain. However, studies have yet to investigate how short-term training on lower-level tonal information translates into a higher-level linguistic domain such as word learning.

The ability to perceive the acoustic distinctions between sounds is a necessary initial step toward being able to utilize them in higher-level linguistic contexts. The importance of acoustic information for speech learning has been demonstrated with both infants (Werker et al., 2002) and adults (Cooper and Wang, 2012; Wong and Perrachione, 2007). Wong and Perrachione (2007) reported that listeners who were highly proficient at identifying pitch patterns, as a result of long-term musical experience, were the most successful learners of vocabulary items distinguished by lexical tones. The authors termed it a “phonetic-phonological-lexical continuity” (p. 566), whereby lower-level auditory abilities that facilitated the perception of phonetic contrasts had a significant impact on a higher-level lexical task. This highlights the
involvement of bottom-up as well as top-down processes in word learning, in that there appears to be a substantive relationship between these different levels of information (i.e., auditory-phonetic-phonological-lexical).

Given that previous research has primarily examined the influence of short-term pitch experience on lower-level abilities such as phoneme identification (Francis et al., 2008; Wang et al., 1999), the current study extends the earlier work of Cooper and Wang (2012) by investigating the effect of short-term pitch experience on higher-level lexical (Cantonese tone-word) learning by training native English speakers on Cantonese tones prior to a tone-word learning program. The results are compared to the English musicians and non-musicians from Cooper and Wang (2012), who completed only the tone-word learning program, to examine how listeners with short-term pitch experience compare to listeners with and without significant (musical) pitch experience. The phonetic-phonological-lexical continuity hypothesis would predict that improving listeners’ tonal awareness and lower-level phonetic abilities through tone training should result in their achieving greater success in a subsequent tone-word learning program relative to non-musician listeners who did not receive any tone training. The non-musicians who receive short-term pitch experience from tone training might also perform at a level on par with musicians who received no lexical tone training but instead benefit from long-term musical pitch experience.

2. Methods

The participants included 32 native English speakers, who had not previously participated in the Cooper and Wang (2012) study, with no prior experience with Cantonese or any other lexical tone language. They self-reported normal hearing and cognitive abilities and were all non-musicians (“N”), defined as having less than 3 yr of musical experience and no experience within the last 5 yr (Wong and Perrachione, 2007). Sixteen of these participants were assigned to a Tone Training group (TT-N; 12 females, 4 males; 

<table>
<thead>
<tr>
<th>Group</th>
<th>Training Type</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-N</td>
<td>None</td>
<td>X</td>
</tr>
<tr>
<td>TT-N</td>
<td>Tone and Word</td>
<td>X</td>
</tr>
<tr>
<td>WO-N</td>
<td>Word-only</td>
<td>—</td>
</tr>
<tr>
<td>WO-M</td>
<td>Word-only</td>
<td>—</td>
</tr>
</tbody>
</table>

The experimental setup for the different groups is illustrated in Table 1. Among the four groups, TT-N participated in all the test and training programs in the following order: (Pre-tone-training) Tone identification (ID) test, Tone training, (post-tone-training) Tone ID test (identical to the pre-tone-training ID test), Tone-word training, and session tests. In order to gauge whether tone training was effective, TT-N’s performance on the two tone ID tests was compared to that of C-N, who also completed these two tests but did not undergo training. WO-N and WO-M participated in one tone ID test followed by the tone-word training program and session tests (Cooper and Wang, 2012).

Table 1. Experimental setup for each group (non-musician control: C-N; tone and word training non-musician: TT-N; word-only non-musician: WO-N, and word-only musician: WO-M). The order of these tasks proceeded from left to right for all groups. The two word-only groups are from Cooper and Wang (2012).
The stimuli used in the pre- and post-test tone ID task were 25 Cantonese words, which included 5 Cantonese consonant-vowel (CV) monosyllables (waj, low, si, pej, fu) with 5 Cantonese tones (High-Level, High-Rising, Low-Falling, Low-Rising, and Low-Level) produced by 2 native Hong Kong Cantonese speakers (1 male, 1 female). Consistent with Cooper and Wang (2012), the Mid-Level tone was not included in any of the tasks to reduce the difficulty level, as it can be easily confused with High-Level, Low-Level, and Low-Rising tones when presented in isolation (Francis et al., 2008). For tone training, 25 different Cantonese words were used, including CV monosyllables (se, jau, tso, seui, ju) produced with the 5 tones by 4 native Cantonese speakers (2 male, 2 female) not used in the tone ID task. Finally, the tone-word training stimuli were the 15 stimuli used in Cooper and Wang (2012), which included 3 CV monosyllables (k“aj, tsou, wu) with 5 tones produced by 4 novel native Cantonese speakers (2 male, 2 female). Each of these 15 tone-word training stimuli was assigned a distinct meaning, as represented by a picture.

In the pre- and post-training tone ID tests, participants were first familiarized with the Cantonese tones by hearing each tone in isolation and viewing its associated tone diagram. Next, participants were familiarized with a five-alternative forced choice ID task (15 trials) with feedback, during which they identified the tone they heard by pressing the number corresponding to the appropriate tone diagram. The format of the main task was identical to the familiarization section; however, feedback was no longer provided. The task consisted of 100 randomized stimuli (5 syllables × 5 tones × 2 speakers × 2 repetitions). For both the TT-N and CN groups, the average time lapse between pre- and post-tests was 9 (range = 7 to 11) days.

The tone training program consisted of three 30-min sessions over the course of an average of 7 days (range = 5 to 8). Each training session was comprised of 300 trials (4 speakers × 5 syllables × 5 tones × 3 repetitions) divided into 6 blocks of 50 trials each. Training was identical in format to the familiarization section of the tone ID task, where listeners identified tones and received feedback on their accuracy.

The tone-word learning task procedures were identical to the program employed in Cooper and Wang (2012). Participants completed 4 days of training, with 2 sessions per day on all but the last day, for a total of 7 sessions. Each word learning session included 5 training blocks, 2 review blocks and a session test. Each training block involved listening to 4 repetitions of 3 words while a picture of the meaning of each word was displayed (12 trials). Participants were then quizzed on the words they learned in each block (12 trials). The review blocks consisted of all 15 words and were first blocked by syllable (Review 1) to draw attention to the tonal distinctions and then randomized (Review 2). Feedback on response accuracy was provided for both the training blocks and the reviews. Each session concluded with a session test, where participants identified 4 repetitions of all 15 tone-words (60 trials) from a choice of 15 options.

3. Results

In order to assess the efficacy of tone training, the proportion of correct responses for the pre- and post-training tone ID tests for TT-N and C-N were submitted to a 2-way analysis of variance (ANOVA) with Group (TT-N, C-N) as a between-subjects factor and Test (pre, post) as repeated measures. The ANOVA yielded a significant main effect of Test \( F(1, 30) = 64.428, p < 0.0001 \), as well as a significant Test × Group interaction \( F(1, 30) = 8.606, p = 0.006 \). Post hoc (Bonferroni) analyses revealed significant improvements from pre- to post-test for both C-N (pre: 54%, post: 62%; \( p < 0.0001 \)), presumably due to practice effect, and TT-N (pre: 58%, post: 74%; \( p < 0.0001 \)). Post hoc (Bonferroni) Group comparisons fixing each level of Test (pre, post) revealed no significant group differences on the pre-test (\( p = 0.525 \)), but significant group differences on the post-test, where TT-N significantly outperformed C-N (\( p = 0.008 \)). These results indicate that listeners who underwent tone training were significantly better at identifying non-native lexical tones by the end of training relative to the control group.
For tone-word training, the average percent correct score was tabulated from each day’s set of session tests to examine overall improvement as well as the time course of learning. A 2-way ANOVA with Training Day (1 to 4) as a repeated measure and Group (TT-N, WO-N, WO-M) as a between-subjects factor (Fig. 1) showed a significant main effect of Day \[F(3,39) = 219.761, \ p < 0.0001\], Group \[F(2,39) = 5.507, \ p = 0.008\] and a Day × Group interaction \[F(6,39) = 3.374, \ p = 0.004\]. Regarding overall improvement after training (by Day 4), Bonferroni-adjusted pairwise comparisons indicated that both TT-N (73%, \ p = 0.008) and WO-M (75%, \ p = 0.007) were significantly more proficient at tone-word ID than WO-N (55%), but did not significantly differ from each other (\ p = 1.0). With respect to their learning trajectories over the word training program, no significant group differences were found on Day 1 (\ p > 0.101). However, by Day 2, WO-M (65%) was significantly more accurate than WO-N (46%, \ p = 0.017), but TT-N (60%) did not significantly differ from WO-N (\ p = 0.080) or WO-M (\ p = 1.0). By Day 3, WO-M (73%) still had significantly higher accuracy scores than WO-N (53%, \ p = 0.007), and while it appears that TT-N (67%) also achieved higher accuracy scores than WO-N, this difference did not reach significance (\ p = 0.057). Bonferroni-adjusted pairwise comparisons of Day fixing each level of Group revealed different learning trajectories: While TT-N made significant improvements on each successive day (\ p = 0.003), both WO-N and WO-M improved up until Day 3 (\ p < 0.008) but failed to significantly improve any further from Day 3 to 4 (\ p = 1.0).

To establish that the two non-musician groups (WO-N, TT-N) did not differ in their initial tone ID abilities, a 1-way ANOVA on the percent correct tone ID scores from their very first tone ID tests (prior to tone training for TT-N, 58%, and prior to word training for WO-N, 48%) revealed no significant group differences \[F(1,28) = 3.086, \ p = 0.09\]. Additionally, as tone ID proficiency was found to play an important role in tone-word learning (Wong and Perrachione, 2007), the three groups’ (TT-N, WO-N, WO-M) tone ID abilities prior to word learning were compared with a 1-way ANOVA on the tone ID scores from the test completed immediately before word learning (but after tone training for TT-N). This showed a significant main effect of Group \[F(2,39) = 24.298, \ p < 0.0001\], with post hoc (Bonferroni) analyses indicating that both TT-N (74%) and WO-M (74%) were significantly more accurate at tone ID than WO-N (48%, \ p < 0.0001). To examine the relationship between tone ID accuracy and word learning, a linear regression model was constructed for all groups, with percent correct word ID scores from the last day as the dependent variable and pre-word learning tone ID scores as the predictor (Fig. 2). Tone ID ability was a significant predictor of word

Fig. 1. Mean percent correct scores for each tone-word identification session test (1–7) organized by day (1–4) and group. Analyses averaged across session tests from each day’s set of training sessions.
learning success \[ R^2 = 0.555, \ F(1,40) = 49.928, \ p < 0.0001 \], indicating that participants with higher tone identification scores before word learning ultimately achieved higher levels of word learning proficiency.

4. Discussion and conclusions

The results of the present study revealed that short-term tone training had a significant impact on the ability to use non-native tones to distinguish word meaning. While all groups obtained similar tone-word ID scores on their first day of word learning (Fig. 1), increasing TT-N’s tonal awareness through tone ID training to a level similar to WO-M resulted in TT-N achieving similar levels of proficiency as WO-M by the end of word learning, with both scoring significantly higher than WO-N. Furthermore, consistent with Wong and Perrachione (2007), tone ID scores were a significant predictor of word ID scores. These results provide further evidence in support of a phonetic-phonological-lexical continuity, whereby improving listeners’ lower-level phonetic abilities (i.e., tone ID), even over the short-term, contributed to greater success in a higher-level linguistic task (i.e., word ID). The present findings highlight the role of bottom-up processes during speech perception and learning, as pointed to in several models of speech perception (McClelland and Elman, 1986; Norris et al., 2000), in that lower-level information can be critical during the early stages of acquisition. Indeed, learners were unable to effectively utilize higher-level contextual information in sentence recognition when lower-level acoustic information was highly degraded (Bradlow and Alexander, 2007).

Musical training or a tone language background have both been found to aid non-native tone-word learning (Cooper and Wang, 2012; Wong and Perrachione, 2007), suggesting that a heightened domain-general attunement to tonal distinctions can be advantageous when learning words distinguished by tonal cues. In the context of these studies, the current findings indicate that for learners without pre-existing tonal awareness from a musical or L1 background, even three tone training sessions can heighten their sensitivity to non-native pitch patterns enough to be highly beneficial in identifying words using such pitch patterns. Tone training drew TT-N’s
attention to the acoustic information relevant for distinguishing the tonal contrasts and enhanced their tonal awareness, in that it improved their ability to isolate and identify components of linguistic units (Wong and Perrachione, 2007). The present results indicate that focusing attention on non-native phonetic information initially can facilitate the encoding of this information into newly forming lexical representations. WO-N needed to construct tone categories while also mapping these features onto lexical representations within the same task. Distinguishing non-native phonetic contrasts has been found to be significantly more challenging in more cognitively demanding tasks such as lexical decision or word learning (Strange and Shafer, 2008), thus compounding the difficulty for WO-N in effectively forming these tone representations. However, by providing tone training prior to word learning, TT-N may have been able to construct more stable or delineated non-native tone categories relative to WO-N, which might have alleviated the cognitive load during word learning.

Musical experience appeared to provide a larger initial gain in lexical identification accuracy relative to TT-N, which may have arisen as a result of WO-M’s many years of pitch experience, as compared to TT-N’s three tone training sessions. However, short-term experience with the tonal contrasts that distinguish these specific lexical items seemed to facilitate steady, significant improvement across all training days. As TT-N significantly improved from Day 3 to 4, unlike the other groups who did not improve, it is conceivable that they would have continued to improve with additional sessions. Further research is necessary to determine whether or not TT-N could achieve higher word identification accuracy as fast as WO-M with a larger amount of initial tone training. Furthermore, future investigations should examine whether or not the word learning trajectories of musicians and tone-trained non-musicians would diverge further with a longer word training program as well as the longevity of the enhanced tonal awareness produced by focused tone training. Brief pitch experience may have beneficial effects for word learning in the short-term, as illustrated by the results in this study; however, future investigations should determine how lasting the effects of this awareness are and whether it can generalize to these tones presented in larger linguistic contexts (e.g., disyllables, phrases).

Acknowledgments

Portions of this work were presented at the 17th International Congress of Phonetic Sciences (2011) in Hong Kong. This study was funded in part by research grants from the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant Nos. 312457-2006 and 312457-2011).

References and links

1To clarify, “tone-words” refer to lexical items minimally contrasted by lexical tones. Tone-word identification refers to identifying the meanings of different lexical items. In contrast, tone identification involves identifying phonemic tone categories.


