Lexical representation of Mandarin tones by non-tonal second-language learners

Jeong-Im Han, and Kimiko Tsukada

Citation: The Journal of the Acoustical Society of America 148, EL46 (2020); doi: 10.1121/10.0001586
View online: https://doi.org/10.1121/10.0001586
View Table of Contents: https://asa.scitation.org/toc/jas/148/1
Published by the Acoustical Society of America

ARTICLES YOU MAY BE INTERESTED IN

The role of F0 and phonation cues in Cantonese low tone perception
The Journal of the Acoustical Society of America 148, EL40 (2020); https://doi.org/10.1121/10.0001523

A typology of laterals in twelve English dialects
The Journal of the Acoustical Society of America 148, EL72 (2020); https://doi.org/10.1121/10.0001587

Effects of consonantal constrictions on voice quality
The Journal of the Acoustical Society of America 148, EL65 (2020); https://doi.org/10.1121/10.0001585

Speech synthesizer produced voices for disabled, including Stephen Hawking
The Journal of the Acoustical Society of America 148, R1 (2020); https://doi.org/10.1121/10.0001490

Sculpting speech from noise, music, and other sources
The Journal of the Acoustical Society of America 148, EL20 (2020); https://doi.org/10.1121/10.0001474

Does a pitch rating method converge on the frequencies within tonal stimuli?
The Journal of the Acoustical Society of America 148, EL99 (2020); https://doi.org/10.1121/10.0001640

ASA
Advance your science and career as a member of the
ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE
Lexical representation of Mandarin tones by non-tonal second-language learners

Jeong-Im Han1,a) and Kimiko Tsukada2

1Department of English, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul, 05029, Korea
2Department of Linguistics, Macquarie University, Macquarie Walk, North Ryde, New South Wales 2109, Australia

jhan@konkuk.ac.kr, kimiko.tsukada@mq.edu.au

Abstract: The present study examined whether second language (L2) learners can lexically encode the tonal contrasts that do not occur in their native language. Advanced Korean learners of Mandarin performed a lexical decision task with medium-term repetition priming. The repeated words were either identical or interchanged with confusable tones. The results showed that advanced L2 learners behaved similarly to the native listeners in processing of L2 tones in favor of the correct lexical items, but still showed differences from native listeners, suggesting that L2 learners may have difficulty in utilizing the L2 sounds for word recognition.

1. Introduction

It is commonly believed that second language (L2) acquisition of lexical tones presents a challenge for learners from nontonal native language (L1) backgrounds, but learners exposed to L2 tonal contrasts are capable of making significant progress in tone identification/discrimination. Crucially for our study, Tsukada and Han (2019) showed that, with the exception of the tone 1–2 pair, advanced Korean learners of Mandarin were as accurate as native listeners in discriminating Mandarin lexical tones. This is noteworthy because, unlike Mandarin, standard Korean spoken in the Seoul/Kyunggi area does not use F0 variations for lexical distinctions at the level of words. Less experienced Korean learners of Mandarin in Tsukada and Han (2019) substantially differed from advanced learners and, to a greater extent, native Mandarin listeners. These results suggest that experienced L2 learners are capable of acquiring nativelike tone perception.

However, these results do not rule out lexical encoding difficulties related to Mandarin tone perception. Several recent studies showed that intermediate or advanced level of L2 listeners had difficulty when Mandarin tone perception involved the lexical encoding, in particular for multisyllabic words, although they successfully identified or discriminated the tones in monosyllabic stimuli (Lee et al., 2009; Hao, 2012; Pelzl et al., 2019). For instance, Pelzl et al. (2019) employed behavioral and event-related potential experiments to examine the perception of Mandarin tones by advanced English learners in isolated syllables, disyllabic words in isolation, and disyllabic words in sentences. They found that English learners excelled at tone identification in isolated syllables, but showed poor performance in both types of disyllabic words, suggesting a genuine difficulty of encoding Mandarin tones lexically.

The purpose of this study is to add to the current understanding of discontinuities between perception and lexical encoding in L2 suprasegmental contrasts that do not occur in their native language. We use a speeded lexical decision task with medium-term repetition priming to examine the robustness of lexical encoding of confusable tonal contrasts in Mandarin by advanced Korean learners. We are most interested in the prime-target pairs where the tones of target words are changed into confusing ones, leading to non-words. These trials are compared to the trials where the prime and target words are identical. The prediction is that if Korean learners of Mandarin can encode Mandarin tone contrasts in their lexicon, they would not show priming effects between the prime words and target nonwords with incorrect tonal specification. In contrast, if they have difficulty in lexical encoding, priming effects would be observed. The target tone pairs are T2-T3, T3-T4, T1-T2, and T2-T4. The choice of these tone pairs is guided by the results of the discrimination task (Tsukada and Han, 2019) such that for advanced level Korean learners, the T2-T3 pair was the hardest for discrimination, while discrimination of the

a)Author to whom correspondence should be addressed.
T3-T4 and T2-T4 pairs was easiest. The T1-T2 pair, which showed intermediate discrimination scores, was further included because this was the only pair with significant difference from the native listeners.

2. Method

2.1 Participants

The participants were 20 Korean learners of Mandarin Chinese (11 females; mean age: 24.35 years, range: 21–31, SD: 2.76). All Korean listeners were advanced level Mandarin learners, having the 5th (n = 3) or 6th (n = 16) grade of Hanyu Shuiping Kaoshi, a standard Chinese language proficiency test, or with experience of living in Mandarin-speaking countries for 12 years (n = 1). The participants had begun learning Mandarin at a mean age of 12.45 years old (range: 1–26, SD: 6.6). Their mean length of learning Mandarin was 11.65 years (range: 3–22, SD: 5.66), and the mean percentage of using Mandarin per day was 23.21% (range: 0–80, SD: 19.47). In addition, 20 native Mandarin speakers (19 females; mean age: 21.5, range: 19–25, SD: 1.82) served as controls, all of whom were college students in Korea. Fourteen of the Mandarin speakers learned English as a regular subject while in China and ten of them learned Korean (mean: 22.5 months, range: 6–36, SD: 11.25).

2.2 Stimuli

A total of 40 monosyllabic words or morphemes were selected for stimulus items, which were distributed across the four distinct tone pair groups, T2-T3, T1-T2, T3-T4, and T2-T4. For each tone pair, nonwords were created by interchanging the tones, such that for the T2-T3 pair, a nonword /kú/ with T3 was created from the real word /kó/ with T2 (“the Buddha”), and a nonword /kú/ with T2 was created from the real word /kú/ with T3 (“bitter”). Five words and five nonwords were allotted for each tone specification in the target pair. The complete list of test words is presented in the supplementary material. In addition, 60 real words (well-formed Mandarin words) and 60 nonwords (real words associated with incorrect, confusable tones) were created as fillers. Given that monosyllabic homophones are prevalent in Mandarin, we tried to choose words which had few homophones. All test and filler words were monosyllabic. Stimuli were recorded naturally in a frame sentence (This word is ___ [zhè gè cíer shì ___] “This word is ___”) by a female Mandarin native speaker (age = 32) in a sound-proof booth at a sampling rate of 44100 Hz with a 16-bit resolution on a mono channel through a Shure microphone (KSM44). From two repetitions, the second recording was used as stimuli of the experiment. In total, 200 items were created (10 words × 4 types of tone pairs + 40 nonwords + 60 filler words + 60 filler nonwords).

2.3 Procedure

The participants were tested individually in a sound-proof booth. The lexical decision task was administered using the software E-PRIME 2.0 PROFESSIONAL. The 200 total items were split into four blocks, separated by a break during the experiment. Each block contained 50 items which included real Mandarin words, Mandarin words with incorrect, confusable tones (nonwords), filler words, and filler nonwords. The target words or tone-changed items were presented 9 to 24 words after the prime words (mean = 14.48, SD = 2.75). The blocks and items within each block were randomized. The participants were first informed that they would hear a series of Mandarin monosyllabic words or nonwords. They were then instructed to decide whether or not each token they heard was a real word of Mandarin by pressing the designated key on the computer keyboard as quickly and accurately as possible. In each block, 500-ms fixation was presented, which was followed by the word item presented auditorily through headphones (Sennheiser 590). If participants did not respond within 3000 ms, the next trial was presented. Before starting the experimental items, there were ten practice trials, which were not included in the actual experiment. The lexical decision task took approximately 20 min to complete. After that, participants completed a background questionnaire and a vocabulary/word familiarity test. For each item (test words and filler words), the participants were asked to write its meanings and familiarity using a seven-point Likert scale (1 = unknown; 7 = very familiar).

3. Results

The total number of trials were 8000 (4000 each for Mandarin and Korean listeners), among which 1600 tokens for the real words and nonwords with tone changes were included for analysis. The data from two Korean learners were not included because their mean accuracy rates were less than 50%. The mean familiarity of the test words was 6.16 (range = 4.85–7; SD = 0.7) for the Korean learners and 6.65 (range = 5.95–7; SD = 0.2) for the native Mandarin speakers, which were not significantly different [F(1, 36) = 3.798, p = 0.06]. Further, the data from no
replies and those with RTs less than 500 ms were removed (three tokens of all responses from Mandarin listeners and eight tokens from Korean learners). In total, 1509 trials were included in the analysis (797 from Mandarin listeners and 712 from Korean learners). Due to high error rates in judgment performance for both groups of participants, only accuracy rates were analyzed for the present analysis.

Figure 1 presents the mean lexical decision accuracy rates (%) from native Mandarin listeners and Korean learners of Mandarin.

The accuracy rates were submitted to a mixed analysis of variance with the participant group (native Mandarin listeners and Korean learners) as a between-subject factor and tone pair type (T2-T3, T1-T2, T3-T4, and T2-T4), and priming condition (repetition and tone change) as within-subject factors. There were main effects of participant group \([F(1, 36) = 23.02, p < 0.0001, \text{partial } \eta^2 = 0.39]\), tone pair type \([F(3, 108) = 3.32, p < 0.05, \text{partial } \eta^2 = 0.084]\), priming condition \([F(1, 36) = 92.36, p < 0.0001, \text{partial } \eta^2 = 0.72]\), and the interaction between participant group and tone pair type \([F(3, 108) = 5.89, p < 0.01, \text{partial } \eta^2 = 0.141]\) and that of the three factors, participant group \(\times\) tone pair type \(\times\) priming condition \([F(3, 108) = 5.15, p < 0.05, \text{partial } \eta^2 = 0.13]\). The results indicate that both native Mandarin listeners and Korean learners performed better on identical (repeated) than test (with tone changes) items, but the native Mandarin listeners showed higher rates in the lexical access than Korean learners. The lexical encoding difficulties for the test items varied according to the tone pair type. Planned pairwise comparisons showed that in the repetition condition, native Mandarin listeners showed significantly higher accuracy rates than Korean learners in T2-T3 and T3-T4 pairs (both \(p < 0.05\)), but the two participant groups did not show significant differences in T1-T2 (\(p = 0.23\)) and T2-T4 (\(p = 0.094\)) pairs. In the tone change condition, the two participant groups showed significant accuracy differences in T2-T3 (\(p < 0.0001\)) and T2-T4 pairs (\(p < 0.05\)), but not in T1-T2 (\(p = 0.068\)) and T3-T4 (\(p = 0.803\)) pairs.

In the separate analysis for the interaction of priming condition and tone pair type in Mandarin and Korean listeners, for Mandarin listeners, there was a significant effect of tone pair type in the tone change condition \([F(3, 57) = 6.308, p < 0.05, \text{partial } \eta^2 = 0.249]\), but not in the repetition condition \([F(3, 57) = 0.106, p = 0.956, \text{partial } \eta^2 = 0.006]\). Similarly, Korean learners showed a significant effect of tone pair type in the tone change condition \([F(3, 51) = 3.997, p < 0.05, \text{partial } \eta^2 = 0.190]\), but not in the repetition condition \([F(3, 51) = 1.882, p = 0.144, \text{partial } \eta^2 = 0.100]\). Simple effects analysis for the tone change conditions further showed that Mandarin listeners showed significant differences between T1-T2 and T2-T3, between T1-T2 and T2-T4, between T3-T4 and T2-T3, and between T2-T4 and T3-T4, while Korean learners showed significant differences between T2-T3 and T2-T4, and between T2-T3 and T3-T4 (all \(p < 0.05\)).

4. Discussion

Using a lexical decision task with medium-term repetition priming, we showed that even advanced Korean learners of Mandarin may have difficulty in lexical encoding of the words with Mandarin tone contrasts. The overall pattern of access to the lexical representations of Mandarin tones appeared to be similar between native Mandarin listeners and advanced Korean learners, but Korean learners’ ability to efficiently process Mandarin tones in favor of the correct lexical
item was still different from that of native listeners. In the T2-T3 pair, which was reported to be hardest in discrimination, Korean learners showed significantly lower accuracy rates than Mandarin listeners in both repetition and tone change conditions, whereas in the T3-T4 and the T2-T4 pairs, which were the easiest pairs in discrimination, Korean learners showed significantly lower accuracy rates than Mandarin listeners in either repetition or tone change condition. Given that most test words showed relatively high familiarity and the lags between the prime and the target were not as long as the ones used in typical long-term repetition priming tasks, the present results indicate that at even advanced proficiency, lexical representations of the words may not be sufficiently detailed and thus the lexical status of the presented tokens cannot be easily detected.

These results are in line with previous research in support of the discontinuities between speech perception and lexical encoding of L2 contrasts such that unlike native listeners, L2 listeners do not properly encode the lexical information of the words containing the confusable categories in L2, and are thereby prone to access the incorrect lexical representation of a confusable competitor, despite their accurate discrimination of the categories with increasing proficiency (Darcy et al., 2013; Nakai et al., 2015). Relative to discrimination, performance in lexical encoding is likely to suffer substantially more as confusability increases, because lexical encoding needs better attentional and general cognitive abilities. For this reason, accurate phonetic categorization would not guarantee successful lexical encoding of a confusable contrast and unreliable lexical encoding would persist into advanced levels of proficiency. In particular, the present results extend the previous results for confusable segmental contrasts to those for suprasegmentals, and thus consistent with the previous results regarding Mandarin tone perception in L2 (e.g., Pelzl et al. (2019)).

With respect to the discontinuities between perception and lexical encoding of lexical tones, Pelzl et al. (2019) pointed out that highly accurate phonetic categorization of Mandarin tones seems to be related to the use of pitch as a prosodic feature in many languages, but even advanced L2 learners may have difficulty in learning how to use pitch as a lexical cue. In fact, native Korean listeners do not use fundamental frequency (F0) variations for distinctions at the level of words, but they do use pitch variations to distinguish discourse meanings at the level of phrase (Jun, 1996). Hence, advanced Korean learners seem to make quick gains in the discrimination of Mandarin tones, but they still have difficulty in using F0 in encoding tones lexically. Direct comparison of the present results and those from the previous study of Tsukada and Han (2019) may not be plausible, but the mean discrimination rates for the four target tone pairs were close to the ceiling (96% for advanced Korean learners and 98% for native Mandarin listeners) in Tsudaka and Han (2019), whereas the lexical decision rates in the present study showed significant differences between these two participant groups (62% for advanced Korean learners and 78% for native Mandarin listeners).

It is noteworthy that even native Mandarin listeners did not show the lexical decision accuracy rates close to the ceiling. This may be related to the fact that tone cues are less informative than segmental cues (Pelzl et al., 2019; Sereno and Lee, 2015). For instance, Sereno and Lee (2015) examined the nature of tonal information in Mandarin spoken word recognition by evaluating the separate contribution of tone and segmental information to word recognition processes. Interestingly, their results showed significant priming effects when both tonal and segmental information overlapped or when only segmental information overlapped, whereas no priming was found when only tones matched. These data clearly indicate that tonal information does not appear to facilitate lexical access. For this reason, even native Mandarin listeners may have difficulties in attending to tonal cues in lexical processing, as shown in the results in Fig. 1.

Another key finding is that Korean learners’ (in)ability to auditorily distinguish the various tone pairs was reflected in their lexical encoding of tone pairs. Namely, their performance was worst on the T2-T3 pair, intermediate for the T1-T2 pair, and best on the T3-T4 and T2-T4 pairs, even though the statistical significance was only observed between the T2-T3 pair and the T3-T4/T2-T4 pairs. These results suggest that lexical encoding abilities drop proportionally as sound perception abilities decrease. Relatedly, for both native and nonnative listeners, the relative difficulty for the T2-T3 tone contrast is presumably based on their acoustic characteristics such as duration and F0 scale and contours (Hao, 2018), which may lead to unsuccessful lexical activation for both L1 and L2 listeners. On the other hand, we acknowledge that lexical decision accuracy for the Mandarin listeners was not directly related to the discrimination accuracy for reasons not yet determined. There might be several reasons for that, one of which could be based on the lexical frequency of the test words across the tone pair types. It might be possible that even native listeners were less successful in the processing of less frequent words. However, close inspection of the test words reveals no main effect of word familiarity in the accuracy differences across the four tone pair types [F(1, 3) = 0.295, p = 0.829], suggesting that word familiarity did
not play a role. Another possible explanation may be based on the use of monosyllabic target words in the present study. Given that the actual Mandarin lexicon is largely disyllabic, the native Mandarin listeners might feel uncomfortable with the processing of monosyllabic tokens. This certainly points to a need for further investigation.

To conclude, the present results suggest that although advanced L2 learners are successful in the discrimination of L2 tones, as lexical processing is needed (i.e., lexical decision), processing the lexical tones in a nativelike fashion is not easy. These results thus support the discontinuities between individual sound perception and lexical encoding in L2 suprasegmental contrasts.

Acknowledgments
We would like to thank Sang-Im Lee-Kim, Hye Won Kwon, and Tae-Hwan Choi, as well as Charles Church and two anonymous reviewers for their contributions to this research.

References and links
1See supplementary material at https://doi.org/10.1121/10.0001586 for experimental stimuli.
2The definition of “word” is not clear in Mandarin Chinese.