INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

# Cognitive factors in Thai-naïve Mandarin speakers' imitation of Thai lexical tones

*Juqiang Chen*[1], *Catherine T. Best*[1,2], *Mark Antoniou*[1]

[1]The MARCS Institute, Western Sydney University, Australia
[2]Haskins Laboratories, New Haven CT, USA
J.Chen2/C.Best/M.Antoniou@westernsydney.edu.au

## Abstract

The present study investigated how cognitive factors, memory load and attention control, affected imitation of Thai tones by Mandarin speakers with no prior Thai experience. Mandarin speakers lengthened the syllable duration, enlarged the F0 excursion and moved some F0 max location earlier compared with the stimuli, even in the immediate imitation condition. Talker variability had a larger impact on imitation than memory load, whereas vowel variability did not have any effect. Perceptual assimilation patterns partially influenced imitation performance, suggesting phonological categorization in imitation and a perception-production link.

**Index Terms**: lexical tones, imitation, cognitive factors, non-native speech contrasts

## 1. Introduction

In order to learn to speak a second language, language learners often rely on imitating the words produced by native speakers. Imitation links speech perception and production in a natural way, providing an excellent opportunity to examine the link between perceiving and producing a non-native language without the need of orthographic knowledge.

It is argued that imitation is constrained by native phonological categories. Native speakers fail to imitate phonetic details of vowel or VOT continua [1-2]. Instead their imitations reflect their native phonological categories. Spanish monolinguals, Spanish speakers of English and English monolinguals were asked to imitate a stop consonant voice onset time (VOT) continuum ranging from /da/ to /ta/ in [2]. Their imitations did not show a linear incremental increase in VOT, but instead formed two or three VOT response categories that matched the native phoneme boundaries. Similarly, Finnish children and adults imitated a /æ/ to /ɑ/ vowel continuum, showing categorical imitation patterns relevant to their native phonological categories [1].

In addition, second language learning research has reported that imitation by second language speakers is modulated by their native language phonology. For example, when asked to imitate eight American English vowels (e.g. /i/, /ɪ/, /e/, /ɛ/, /æ/, /ʌ/, /ɑ/, /u/), native Mandarin speakers showed influence of their native language [3]: /ɛ/, /æ/ with no Mandarin counterparts were imitated worse than /i/ and /u/ with counterparts in Mandarin. These findings support well-established second language speech learning theories, such as the Speech Learning Model (SLM) [4] and the Perceptual Assimilation Model (PAM) [5]. Both models claim that attunement to the native language will cause listeners' inability to discern the phonetic differences between pairs of sounds in the L2, or between L2 and L1 sounds via "equivalence classification" (SLM) and "perceptual assimialtion" (PAM). In addition, without accurate perceptual "targets" to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will be inaccurate [4].

However, some researchers argue against imitation as always being phonologically constrained [6]. English speakers identified and imitated Mandarin tones whereas Korean speakers identified and imitated English consonants. Researchers found that imitation was generally more accurate than the identification. Therefore, they inferred that imitation was not always constrained by native phonology as in identification and that L2 imitation may bypass some aspects of phonological encoding. Furthermore, they proposed that participants can operate on a phonetic mode of processing without any access to native phonology. In [7], native speakers of Polish imitated English unreleased plosives in two-stop sequences in two imitation conditions. In one condition, they imitated the stimuli immediately and, in another condition, they read a digit after the auditory input before imitating. Performance was native-like in the first condition and it was significantly impeded in the second. This suggests that native language constraints on imitation of non-native phones is modulated by the mode of imitation. According to the Automatic Selective Perception (ASP) model, the change between a phonetic and phonological mode in perception is modulated by cognitive load [8]. In this paper, we extend ASP to imitation to make predictions about how cognitive factors may influence imitation performance.

Very few studies have investigated non-native tone imitation by tone language speakers. Even fewer tone imitation studies have considered cognitive factors. Therefore, the present study examines how cognitive load factors affect the imitation of Thai tones by Mandarin-native speakers with no prior experience with Thai. Thai and Mandarin differ in the number of tones and types of tones in their native inventories. We used Chao values [9] to provide a priori phonetic description of the tones in each language. In Chao notation, F0 height at tone onset and offset is referenced by numbers 1-5 ranging from low to high. Thai has three level tones (characterized as high-level T45, mid-level T33, low-level T21) and two contour tones (rising T315 and falling T241) [10]. Mandarin has four tones: a level tone M55; a rising tone M35; a falling-rising tone M214; and a falling tone M51[11]. Pervious perception work [12] has indicated that Thai-naïve Mandarin listeners assimilated T45 and T315 into a single Mandarin tone category, M35. T33 and T21 were categorized as M55 and M214 respectively. T241 was split between M55 and M51, thus uncategorized.

Two cognitive factors, namely memory load and attention control demand, were systematically manipulated in the study. Memory load is the capacity to hold a rapidly decaying

http://dx.doi.org/10.21437/Interspeech.2019-1403

memory for a limited period of time [13]. It can affect imitation because the auditory memory used in the phonetic mode decays quicker than the memory of more abstract, "encoded" phonological categories. Thus the longer participants are asked to hold the tone in memory, the greater the auditory memory decays, and the more they will have to rely on their longer-lasting phonological memory for imitation. Attention control is the capacity to efficiently allocate attention between task-relevant and irrelevant information [14]. The more complex the stimuli, the higher the demand on attention control, which can affect imitation because participants have to allocate more cognitive resources to process the stimuli to extract tone-related information.

We hypothesize that imitation will be more accurate (i.e., less deviant from the stimulus) when memory load is low because the phonetic information is still available in short memory. Moreover, when acoustic complexity of the stimuli within a block is low, e.g. from one speaker and of one vowel, imitation will be better because participants can attend to tone-related phonetic details. Furthermore, the speaker's native language will interact with the cognitive factors, specifically, since participants in this study are naïve listeners, they have no L2 phonological system to use, imitation will be more constrained by L1 phonology and more L1 accented when memory load is high and stimuli within one block are acoustically complex.

# 2. Experiment

## 2.1. Method

### *2.1.1. Participants*

28 native speakers of Mandarin participated in the experiments, divided into two groups for each memory load condition (low load: $M_{\text{age}}$ = 24 years, *SD* = 4; 8 females; high load: $M_{\text{age}}$ = 25 years, *SD* = 6; 10 females). Participants completed a background questionnaire before the test. All had normal hearing and none had experience with Thai or more than two years of formal musical training because musical training can facilitate tone perception and imitation [15].

### *2.1.2. Stimulus materials*

Two syllables (/ma/, /mi/) were chosen for the target stimuli because they are real words in Thai and Mandarin. The target Thai syllables were each recorded several times as produced by two female native Thai speakers who had no experience with any other tone languages. Two tokens of each target item judged to be correct and natural-sounding to a third Thai speaker were used in the imitation study.

### *2.1.3. Procedure*

Memory load was operationalized as the time between the end of the stimuli and the signal for participants to imitate (imitation interval). In the low memory load condition, a message "Imitate now!" was shown 500 ms from the offset of the stimulus to let participants start imitating. In the high memory load condition, it was shown 2000 ms after the offset of the stimulus. In both conditions, participants had 3 s (timeout) to imitate and the inter-trial interval is 1 s.

Attention control was operationalized as within-block talker variability (same vs. different) and vowel variability (same vs. different vowels: /ma/, /mi/). Participants were instructed to imitate stimuli as faithfully as possible after they heard the auditory stimulus. Before the test session, participants completed 10 practice trials. Each participant had 160 trials (2 syllables × 5 tones × 8 conditions × 2 repeats) in total.

Participants were tested individually in testing booths (at Western Sydney University, or UNSW). Stimuli were presented on a Dell Latitude 7280 laptop running E-Prime Professional 2 via Sennheiser HD 280 Pro headphones at 72 dB SPL. Participants' responses were recorded with a portable digital speech recorder (ZOOM H4n) with 41 kHz sampling rate and 16-bit stereo format.

### *2.1.4. Data analysis*

Average pitch, direction, length, extreme point and slope have been reported to be the primary factors affecting the perception of lexical tones [16]. While tone contour direction and slope can be compared via several statistical modelling methods, such as growth curve analysis [17], generalized additive mixed modeling [18] and functional data analysis [19], due to the limited space, this paper focuses on comparing discrete measures that capture features like average pitch, length and extreme points. ProsodyPro [20], a Praat script, was used to extract several acoustic measures from the pitch contours of the Thai stimuli and their imitations: syllable duration, F0 mean, time-normalized 10 points of F0, and F0max location (relative to the syllable duration). F0 maximum to minimum excursion (F0 excursion in short) was calculated by measuring the range of F0 between 10% to 90% of the syllable length (the most stable part). Theoretically, F0 means indicate overall pitch for the three phonologically level tones: T33, T21, T45. F0 excursion could distinguish level tones from contour tones, T241 and T315 which should differ in having different F0max locations. Statistically, in a PCA analysis of lexical tones [21], these acoustic measures outweighed other measures in differentiating Thai, Mandarin, Southern and Northern Vietnamese tones. In order to make F0 means comparable across different speakers, we did Lobanov normalization to F0 means and calculated F0 excursion based on Lobanov-normalized F0 means [22]. It should be noted that the Lobanov-normalized F0 mean values reflect how much an F0 mean for a tone varies from the F0 mean of the speaker.

## 2.2. Results

2.2.1 Acoustic comparison of Thai tones target stimuli

First, we measured syllable duration, Lobanov-normalized mean F0, F0 excursion, F0max location (in Table 1) to examine how these measures contribute to distinguishing Thai tone target stimuli.

Four linear mixed-effects models were built with the four measures as dependent variables respectively and tone types as the fixed-effects factor and participants and vowels as random-effects factors. To calculate the *p*-values for the fixed effects (tone types), we used the Kenward-Roger approximation to the degrees of freedom, as recommended by [23], and the *Anova* function from the *car* package in R, with test specified as "F". Significant main effects of tone types as a fixed factor were found for all four measures: syllable duration, $F(4, 33) = 3.10, p = .03$; mean F0, $F(4, 33) = 22.59, p < .001$, F0max-min excursion size, $F(4, 33) = 12.60, p < .001$, F0max location, $F(4, 33) = 72.63, p < .001$. This indicates that the selected four measures distinguish the five Thai tones.

Moreover, we conducted multiple comparisons to test how different measures distinguish Thai tones with the R-package *lsmeans*. *P*-values smaller than .05 were considered significant. T241 was significantly shorter in syllable duration than T315 and T45, whereas differences in syllable duration among other Thai tones were not significant. As for F0 mean, three phonologically level tones, T33, T21, T45 were distinct from each other. All other tone pairs were significantly different in F0 mean, except for T241-45 and T315-33. T241 had a significantly larger F0 excursion than all other Thai tones whereas T33 had a significantly smaller F0 excursion than other Thai tones. Both T21 and T33 showed no difference in F0 maximum location. T241 had the maximum F0 in the middle of the syllable while both T45 and T315 had it at syllable offset.

Table 1: *Acoustic measures (means) for the Thai target stimuli*[a]

| Thai tones | Duration (ms) | F0_mean | F0 excursion | maxF0_loc (%) |
|---|---|---|---|---|
| T21 | 597 | -0.11 | 0.19 | 13 |
| T241 | 548 | 0.07 | 0.27 | 38 |
| T315 | 625 | -0.02 | 0.16 | 98 |
| T33 | 596 | -0.03 | 0.08 | 29 |
| T45 | 612 | 0.08 | 0.19 | 89 |

[a]Note: F0 mean and F0 excursion are normalized using formula in [22].

### 2.2.2 Deviation of the imitated Thai tones from the targets

4640 raw imitated tones were collected and 179 were removed because participants started imitating before they had been instructed to do so. To obtain difference scores, we subtracted stimuli data from the imitation data for each acoustic measures: syllable duration, Lobanov-normalized mean F0, F0 excursion, F0max location. The resulting difference scores were selected as dependent variables and each was fitted with a linear mixed-effects model. Memory load (low vs. high), talker variability (same vs. different), vowel variability (same vs. different) and tone types (five Thai targets) were used as fixed factors and participant and imitated vowel (high vowel /i/ and low vowel /a/) were random factors (intercept). Four models were built to test all possible main effects and interactions.

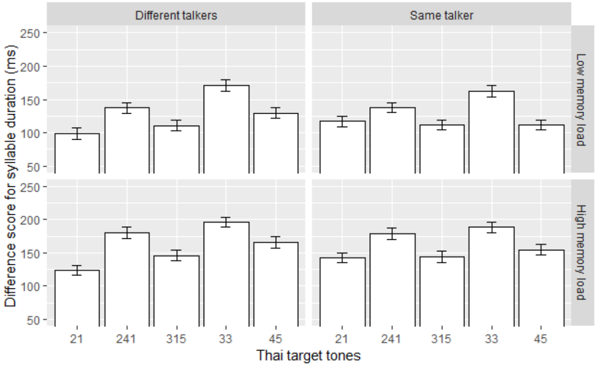

Figure 1: *Difference scores for syllable durations under different cognitive load conditions*

Syllable duration of the imitated Thai tones is shown in Figure 1. We found a significant main effect for tone types, $F(4, 4394) = 69.74$, $p < .001$ and a significant interaction between talker variability and tone types, $F(4, 4394) = 4.58$, $p < .001$. No main effects or interactions were found for vowel variability. We ran multiple comparisons to test the pairwise differences among tone types and the interaction. Positive difference scores indicate that Mandarin speakers lengthened the syllables relative to the original target durations. T33 was significantly lengthened as compared to all other Thai tones. T241 was significantly lengthened as compared to T45, T21, T315, and T45 imitations were significantly longer than the targets for T21 and T315. There was a significant effect of talker variability for T21, $\beta = -18.50$, $SE = 5.71$, $t(4394) = -3.241$, $p = .03$, but not for any other Thai target tones.

Difference scores for Lobanov-normalized F0 mean of the imitated Thai tones are shown in Figure 2. Normalized F0 showed significant main effects of talker variability, $F(1, 4394) = 9.50$, $p = .002$, and tone types, $F(4, 4394) = 54.31$, $p < .001$ and significant interactions between talker variability and tone types, $F(4, 4394) = 11.01$, $p < .001$, and between imitation interval and tone types, $F(4, 4394) = 13.06$, $p < .001$.

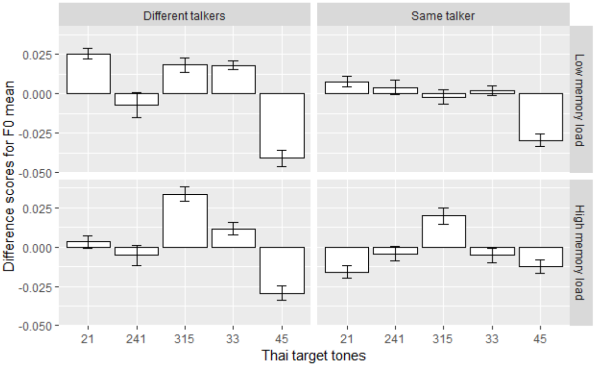

Figure 2: *Difference scores for Lobanov-normalized F0 mean under different cognitive load conditions*

We ran multiple comparisons to test the pairwise differences among tone types and to break down interactions between tone types and memory loads, and between talker variability and tone types. The imitations of both T45 and T315 had significantly larger difference scores compared with other tone imitations. Memory load significantly affected T21, $\beta = 0.02$, $SE = 0.0046$, $t(592) = 4.90$, $p < .001$, T315, $\beta = -0.018$, $SE = 0.0046$, $t(590) = -4.01$, $p = .002$, and marginally significant for T45, $\beta = -0.014$, $SE = 0.0046$, $t(596) = -3.09$, $p = .06$. Talker variability had a significant effect on T21, $\beta = -0.0185$, $SE = 0.0046$, $t(4394) = 3.99$, $p = .0027$, T315, $\beta = -0.017$, $SE = 0.0046$, $t(4394) = 3.70$, $p = .008$, T33, $\beta = -0.016$, $SE = 0.0046$, $t(4394) = 3.57$, $p = .01$ and had a marginal significant effect of T45, $\beta = -0.014$, $SE = 0.0046$, $t(4394) = -3.06$, $p = .06$.

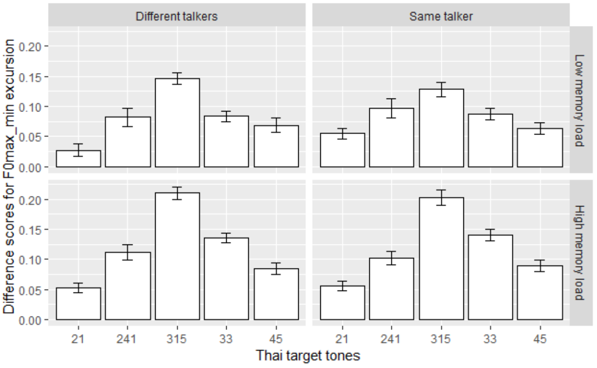

Figure 3: *Difference scores for F0 excursion under different cognitive load conditions.*

For F0 excursion (in Figure 3), we found significant main effects of tone types, $F(4, 4394) = 86.03$, $p < .001$ and a significant interaction between memory load and tone types, $F(4, 4394) = 6.29$, $p < .001$. The difference scores for all tones were positive, suggesting that imitated tones had larger excursion size than the original stimuli. We ran multiple comparisons to test the pairwise differences among tone types and the interaction. T315 had significantly larger difference scores than other tones while T21 had a significantly smaller difference score than other tones. Both T241 and T33 had significantly larger difference scores than T45. Memory load did not significantly affect excursion difference scores for the same tone targets.

F0 max location ratio (Figure 4) showed main effects of talker variability $F(1, 4394) = 7.55$, $p = .006$, and tone types, $F(4, 4394) = 95.46$, $p < .001$, and significant interactions between talker variability and tone types, $F(4, 4394) = 40.95$, $p < .001$ and memory load and tone types, $F(4, 4394) = 5.29$, $p < .001$. The negative difference scores in F0 maximum location indicates earlier position of maximum location in the syllable for imitation than stimuli. We ran multiple comparisons to test the pairwise differences among tone types and the interactions. The difference scores were larger in both T241 and T315 as compared with other tones. Talker variability had significant effects on T21, $\beta = -0.096$, $SE = 0.01614$, $t(4394) = -5.95$, $p < .0001$, T241, $\beta = -0.1224$, $SE = 0.0161$, $t(4394) = -7.58$, $p < .0001$, T33, $\beta = -0.1421$, $SE = 0.0161$, $t(4394) = 8.78$, $p < .0001$. Memory load did not significantly affect F0 max location difference scores for the same tone targets.

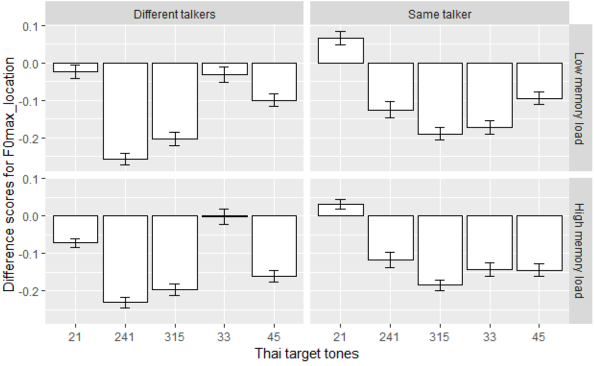

Figure 4: *Difference scores for F0 maximum location ratio under different cognitive load conditions*

## 3. Discussion

First of all, we found a significant main effect of tone types for all four acoustic difference measures. In other words, the deviation in imitation from the target stimuli in all four measures varied from one tone to another. T21 was the best imitated tone, showing low difference scores in most cases but it was susceptible to memory load and talker variability. The imitations of T241 had significant larger difference scores in syllable duration and F0 excursion and F0max, suggesting that T241 was difficult for Mandarin participants. Given that T241 was not categorized as any Mandarin tone categories [12], the L1 phonological influence is smaller than categorized tones. Thus the observed difficulties in imitation lie more in encoding and producing the Thai targets. T33 was significantly lengthened and was imitated with a larger F0 excursion as compared to the target stimuli. T315 was imitated with higher F0, whereas T45 with lower F0 than the original stimuli. Given that T315 is lower than T45, this means that the two tones were produced with very similar F0 which is in line with the observation that T315 and T45 were both perceptual assimilated into M35 [12]. To sum up, Mandarin speakers lengthened the syllable duration to match the longer stimulus syllables, but they overdid the lengthening. F0 excursion was enlarged in the imitations, suggesting that Mandarin speakers attempted to follow the pitch change of Thai, but over-exaggerated the change. Mandarin speakers tended to delay their F0 peak for T241, T315 and T45, relative to the target stimuli.

Second, some cognitive factors affected imitation performance. Talker variability had significant main effects on two of our four measures, namely F0 mean and F0 max location. It was also involved in interactions with tone types for syllable duration, F0 mean and F0 max location measures. This supports our hypothesis that processing task-irrelevant information increases demands on attention control, leading to poorer performance when imitating.

Memory load did not show a main effect for any of the acoustic measures, but it interacted with tone types for F0 mean, F0 excursion and F0 max location. In multiple comparison tests, for the same tone, memory load modulated F0 mean for T21, T315 and T45. However, for F0 excursion and F0 max location, we did not find any significant effect of memory load for the same tone types. Therefore, memory load did not drastically change imitation performance. The difference scores from the target stimuli in the low memory load condition suggest that participants did not imitate the phonetic details of the target tones correctly. Previous research on imitation of consonant length also showed little effect of memory load [24]. The less than expected effect of memory load could be because when waiting for imitation in the long interval condition, participants rehearsed internally, which reduced the decay of phonetic details perceived from the stimuli, thus diminishing the difference between two memory load conditions. Studies that asked participants to do other tasks while waiting have reported stronger effect of memory load [7].

Vowel variability within a block did not have any main effects or interactions on the imitation of Thai tones for any of the acoustic measures. Unlike talker variability which affects pitch more drastically and requires listeners to adapt to a change in talkers, vowel quality is more intrinsic to lexical tones for tone language speakers. Therefore, processing vowel variability may require less cognitive effort than resolving talker variability.

## 4. Conclusions

The present study examined how cognitive factors, memory load and talker variability, affected imitation of Thai tones by Mandarin speakers with no experience with Thai. Mandarin speakers lengthened the syllable duration, enlarged the F0 excursion and moved F0 max location earlier, even in the immediate imitation condition. Talker variability affected imitation most and memory load altered imitation performance to a lesser degree. Vowel variability did not have any effect on imitation. Perceptually uncategorized tones are difficult for naïve speakers to imitate and when two tones are assimilated into a single category, the imitation of these two tones resembles each other. These results have implication for theories of non-native speech perception and production (SLM, PAM) as well as pedagogical implications for second language lexical tone training.

# 5. References

[1] L. Alivuotila, J. Hakokari, J. Savela, R.-P. Happonen, and O. Aaltonen, “Perception and imitation of Finnish open vowels among children, naïve adults, and trained phoneticians,” presented at the Proceedings of the 16th International Congress of Phonetic Sciences, 2007, pp. 361–364.

[2] J. E. Flege and W. Eefting, “Imitation of a VOT continuum by native speakers of English and Spanish: evidence for phonetic category formation,” *J. Acoust. Soc. Am.*, vol. 83, no. 2, pp. 729–740, Feb. 1988.

[3] G. Jia, W. Strange, Y. Wu, J. Collado, and Q. Guan, “Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure,” *The Journal of the Acoustical Society of America*, vol. 119, no. 2, pp. 1118–1130, Jan. 2006.

[4] J. E. Flege, “Second-language speech learning: Theory, findings, and problems,” in *Speech perception and linguistic experience: Issues in cross-language research*, W. Strange, Ed. 1995, pp. 229–273.

[5] C. T. Best, “A direct realist view of cross-language speech perception.,” in *Speech perception and linguistic experience: Issues in cross-language research*, W. Strange, Ed. Timonium, MD: York Press, 1995, pp. 171–204.

[6] Y.-C. Hao and K. de Jong, “Imitation of second language sounds in relation to L2 perception and production,” *Journal of Phonetics*, vol. 54, pp. 151–168, Jan. 2016.

[7] A. Rojczyk, A. Porzuczek, and M. Bergier, “Immediate and Distracted Imitation in Second-Language Speech: Unreleased Plosives in English,” May 2013.

[8] W. Strange, “Automatic selective perception (ASP) of first and second language speech: A working model,” *Journal of Phonetics*, vol. 39, no. 4, pp. 456–466, Oct. 2011.

[9] Chao. Y.R., “A system of tone-letters,” *Le Maitre Phonetique*, vol. 45, pp. 24–27, 1930.

[10] A. Reid *et al.*, “Perceptual assimilation of lexical tone: The roles of language experience and visual information,” *Atten. Percept. Psychophys.*, vol. 77, no. 2, pp. 571–591, Feb. 2015.

[11] M. Yip, *Tone*. Cambridge: Cambridge University Press, 2002.

[12] J. Chen, C. T. Best, M. Antoniou, and B. Kasisopa, “Cross-language categorisation of monosyllabic Thai tones by Mandarin and Vietnamese speakers: L1 phonological and phonetic influences,” presented at the Proceedings of the Seventeenth Australasian International Conference on Speech Science and Technology, 2018, pp. 168–172.

[13] A. Baddeley and B. A. Wilson, “Prose recall and amnesia: implications for the structure of working memory,” *Neuropsychologia*, vol. 40, no. 10, pp. 1737–1743, 2002.

[14] T. Isaacs and P. Trofimovich, “Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of second language speech,” *Applied Psycholinguistics*, vol. 32, no. 1, pp. 113–140, Jan. 2011.

[15] T. L. Gottfried, A. M. Staby, and C. J. Ziemer, “Musical experience and Mandarin tone discrimination and imitation,” *The Journal of the Acoustical Society of America*, vol. 115, no. 5, pp. 2545–2545, Apr. 2004.

[16] J. T. Gandour, “The perception of tone,” in *Tone: A linguistic survey*, Academic Press, 1978, pp. 41–76.

[17] P. Tang, I. Yuen, N. X. Rattanasone, L. Gao, and K. Demuth, “Acquisition of weak syllables in tonal languages: acoustic evidence from neutral tone in Mandarin Chinese,” *Journal of Child Language*, vol. 46, no. 1, pp. 24–50, Jan. 2019.

[18] M. Wieling, “Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English,” *Journal of Phonetics*, vol. 70, pp. 86–116, Sep. 2018.

[19] M. Gubian, F. Torreira, and L. Boves, “Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts,” *Journal of Phonetics*, vol. 49, pp. 16–40, Mar. 2015.

[20] Y. Xu, “ProsodyPro—A tool for large-scale systematic prosody analysis,” 2013.

[21] J. Chen, C. T. Best, M. Antoniou, and B. Kasisopa, “Mapping and comparing East and Southeast Asian language tones,” presented at the Australia Linguistic Society annual conference, Adelaide, 2018.

[22] B. M. Lobanov, “Classification of Russian Vowels Spoken by Different Speakers,” *The Journal of the Acoustical Society of America*, vol. 49, no. 2B, pp. 606–608, Feb. 1971.

[23] U. Halekoh and S. Hojsgaard, “A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models–the R package pbkrtest,” *Journal of Statistical Software*, vol. 59, no. 9, pp. 1–30, 2014.

[24] Y. Asano and B. Braun, “Does speech production in L2 require access to phonological representations?,” presented at the Proceedings of the International Conference on Speech Prosody, 2016, vol. 2016-January, pp. 237–241.