



# Cross-language Lexical-tone Identification

Jennifer A. Alexander<sup>1,2</sup>, Yue Wang<sup>1</sup>

<sup>1</sup> Department of Linguistics, Simon Fraser University, Canada

<sup>2</sup> Department of Linguistics, Northwestern University, USA

jennifer.alexander@northwestern.edu, yuew@sfu.ca

## Abstract

We examine how native lexical-tone experience influences identification of novel tone. Cantonese, Thai, Mandarin, and Yoruba listeners identified CV syllables bearing the six phonemic Cantonese tones. Overall tone accuracy scores and tone-error patterns were assessed. Consistent with previous reports, native listeners' perception reflected effects of ongoing tonal mergers and acoustic similarities incurred by a crowded lower tone space [27]. Non-native listeners' perception appeared to be influenced by L1 auditory experience, as they attended to acoustic and phonetic cues relevant to the phonological and phonetic properties of L1 tone categories. [Support: NSF grant 0965227 to J.A.A.]

**Index Terms:** cross-language, lexical tone, identification, perception, Cantonese, Thai, Mandarin, Yoruba, PAM

## 1. Introduction

During perception of novel speech sounds, listeners attend to gradient (non-contrastive) phonetic information, as well as acoustic cues that signal a phonological contrast in the native language (L1) [1-3]. The Perceptual Assimilation Model (PAM) attempts to account for this and the observation that all non-native sound contrasts are not equally well- or poorly-perceived [*ibid.*]. Much of the work in this area examines perception of vowels [4-5] and consonants [6-7]; studies on cross-language suprasegmental-contrast perception focus primarily on quantity, stress, and sentential prosody [8-9]. Yet, perception of non-native tone is likely constrained by experience with the phonetic and phonological properties of the L1 tone system as well. Prior research examines tone perception in non-tone listeners or in 1-2 tonal-L1 groups; most of these test perception of Mandarin or Thai tones [10-16]. We seek to replicate and extend this line of study, testing native- and non-native Cantonese tone identification in listeners from languages with different tone-inventories: Cantonese, Thai, Mandarin, and Yoruba. We aim to provide a uniquely comprehensive and generalizable view of the effect of tone-language experience on perception of non-native tone. Results are discussed with reference to the Perceptual Assimilation Model for Suprasegmentals (PAM-S) [10-11].

### 1.1. Lexical-tone languages

In tone languages, F0 (pitch)-contour and -height contrasts over a tone-bearing unit signal word meaning. Hong Kong Cantonese has three level tones differentiated acoustically and perceptually by relative pitch height (high (CHL), mid (CML) and low (CLL)) [17], and three contour tones differentiated by direction and magnitude of F0 change (low/mid falling (CMF), mid-rising (CMR), and low-rising (CLR)) [18]. Bangkok Thai has three level tones (high (THL), low (TLL), and mid/neutral (TN)); each is actually contoured) and two contour tones (falling (TF) and rising (TR)) [19]. Beijing

Mandarin has one (high-)level tone (MHL) and three contour tones (rising (MR), falling-rising (MFR), and falling (MF)) [20]. Yoruba has high (YHL), mid (YML), and low (YLL) level tones [21-22]; isolated YLL falls. Fig. 1 shows pitch tracks of the languages' tones (F0 in semitones/ST) of citation-form tones on [gi] produced by 1 male native speaker.

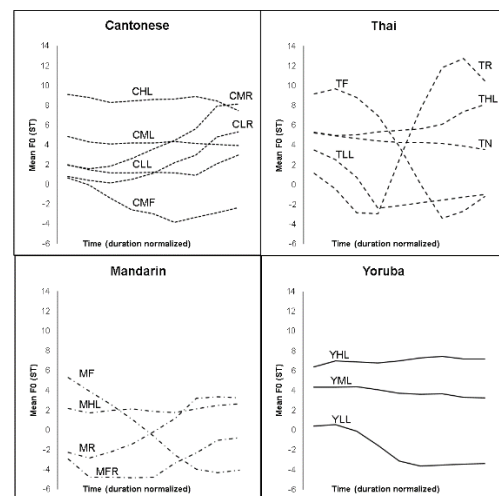


Figure 1: F0 tracks of the tones of Cantonese (panel A), Thai (B), Mandarin (C), and Yoruba (D)

### 1.2. Perceptual processes and the current study

Phonetic and phonological assimilation processes underlie perception of segments and suprasegmentals alike. *Phonological* assimilation is a high-level process during which a listener maps a non-native speech sound to a native (L1) phoneme with a similar lexical function [23]. Dutch and German listeners, for instance, perceive American English /ɪ/-/i/ categorically, presumably due to their phonetic similarity to L1 phonemic contrasts [7]. But Japanese listeners confuse and assimilate English /ɪ/-/i/ to Japanese /i/ [24], as these sounds are non-contrastive in Japanese. *Phonetic* assimilation occurs when attendance to gradient acoustic-phonetic cues results in perceptual equivalence of two or more non-native sounds. For American English listeners, Zulu /kʰ/ and /k'/ assimilate to English [kʰ]; the former is a close perceptual match to English [kʰ], and since English lacks ejectives, /k'/ is heard as a deviant version of [kʰ] [25]. As such, listeners perceive these non-native phonemes as non-contrastive.

As in non-native segment perception, non-native tone perception involves attendance to acoustic cues relevant to the phonological and phonetic properties of the L1 tone system. The PAM-S proposes that tonal and non-tonal listeners both assimilate non-native prosodic contrasts to L1 prosodic categories [10-11]. The model predicts that a novel sound will be *categorized* if perceived as belonging to a single L1 prosodic category, *uncategorized* if it is perceived as phonetically similar to multiple L1 prosodic categories or

unknown/not clearly assimilable to any. It posits five patterns of assimilation. In *two-category (TC)* assimilation, two non-native prosodic categories are assimilated to two different L1 prosodic categories; discrimination of these sounds is excellent. *Single-category (SC)* contrasts – two non-native prosodic categories that are equivalently good or poor examples of a single L1 prosodic category – are poorly discriminated. When two non-native sounds are assimilated to a single L1 category but differ in goodness-of-fit, good to moderate discrimination results; this is a *category-goodness (CG)* difference. When the sounds are *uncategorized*, discriminability depends on the similarity between the sounds and one or more native categories.

As an example of the above, Mandarin listeners (henceforth *Mand. listeners*) attend primarily to F0 direction during native-tone perception [26]. In turn, naïve Mand. listeners have been found to accurately identify CHL, CMR, and CML, arguably because they bear F0 trajectories like those of the MHL, MR, and MFR, respectively [16]. This result exemplifies *phonological* assimilation in novel- (Cantonese) tone perception as the listeners assimilated novel tones to functionally-similar L1 tone categories. In PAM-S terms, these are instances of *TC* assimilation. However, Mand. listeners are insensitive to the height differences between CMR and CLR. They also confuse CHL, CML, and CLL [16, 27]. This illustrates *phonetic* assimilation in tone perception, since for these listeners, such pitch-height differences reflect within-category variation, rather than category distinctions, in the L1 [28-29]. These would be *SC* contrasts. We anticipate replicating these results. Thai listeners attend to numerous phonetic cues for native tone identification, including endpoint F0, duration, amplitude, and turning point [30-32], but can identify L1 tones on the basis of F0 slope and direction alone [33]. F0 height also functions as a secondary cue for level tones [32]. As would be expected, naïve Thai listeners have been found to attend to F0 height and direction to assimilate novel Mandarin tones to the closest L1 tone [13]. For instance, THL assimilates to MHL, perhaps because these tones have similar offglide F0 values. This constitutes phonological, *TC*, assimilation. Our Thai listeners might perform similarly in response to Cantonese tones. But we also expect confusion patterns that stem from the acoustic-phonetic dissimilarities between similarly-labeled L1 and Cantonese tones. Listeners might, e.g., correctly identify CML, mapping it to TN per their F0 height and direction similarities (*phonological, TC*, assimilation). But they might map both CLL and CML to TML since Thai has only one truly level tone to serve as a level-tone target (*CG* assimilation of CLL and CML to TML). Likewise, Thai listeners might find the relatively static CHL tone *uncategorizable*, since it differs so profoundly in F0 height and direction from the so-called high-level tone of Thai. Yoruba tones are distinguished primarily by relative F0 height in isolation [21]. Listeners also attend to F0 direction for YLL identification, as YLL falls in isolation and in utterance-final position [34]. Slope and offglide F0 are also important but are assigned less weight [*ibid.*]. Like the Thai listeners, the Yoruba listeners (henceforth *Yor. listeners*) might map both CML and CLL to YML, since these tones have similar offglide F0 and mean F0 values (YML’s height being between the two Cantonese tones’ heights). This would reflect *phonetic, SC* or *CG*, assimilation, depending on the extent to which each novel tone was confusable with the others. Crucially, these tones would not be expected to map to YLL because, unlike YLL, they don’t fall.

Perceptual accuracy is also potentially inversely related to the tones’ proximity in the acoustic-perceptual tone space. The CMR-CLR contrast is confusable even for adult native

listeners [27], as are the CML-CLL and CMF-CLL contrasts. This might be because the tones in each pair (a) differ only subtly in both F0 direction and overall mean F0; (b) occupy a small, sometimes-overlapping, overall pitch range; and (c) are merging, thereby becoming less acoustically and perceptually distinct [*ibid.*]. Lack of phonetic context can exacerbate these perceptual issues. Our stimuli are presented in isolation, so we expect the native listeners to make these types of confusions.

## 2. Methods

### 2.1. Participants

Participants were 55 adult native speakers of Cantonese (n=15, 8 female; mean age (M)=22.6 years), Thai (n=13, 8 female; M=27 years), Mandarin (n=15, 13 female; M=21 years), and Yoruba (n=12, 3 female; M=22 years) from the Vancouver, B.C., area. All were literate in English and their L1; none reported any speech, reading, hearing, or neurological problems. Place-of-origin was controlled to the extent possible to minimize dialect variation. Yoruba participants were from Nigeria (8 from Lagos); Cantonese, Thai, and Mandarin participants were raised within 185 miles of, respectively, Hong Kong or Guangzhou; Bangkok; or Beijing or Taipei. Participants had little-to-no music experience (max 7 years, n = 1; mode 0-2 years, n = 38), important because music training is shown to affect perception and production of non-native tones [35]. Participants were also fluent in no tone language but the L1, thus avoiding perceptual interference from other tone languages. Participants were compensated.

### 2.2. Materials

Stimuli were 108 tokens of Cantonese-tone-bearing isolated monosyllables: *ba, bi, bu, da, di, du* were each produced with the six Cantonese tones, by one male talker, three times. Most were nonwords. The stimuli were sampled at 44.1 kHz, normalized for RMS amplitude at 65 dB, and normalized for average duration at 0.42 s. These steps were critical as syllable duration and amplitude envelope are cues for tone perception [36]. These syllables are phonotactically legal in all four languages, so listeners were expected to be able to focus on the tones rather than the segments or syllable structure. To confirm the tones’ intelligibility, four native Cantonese informants rated their accuracy on a whole-number scale of 1 (poor) to 5 (perfect). Mean and median tonal accuracy, collapsed across all syllables, were 4.5 and 4.6, respectively.

### 2.3. Procedure

Participants performed a 6AFC speeded tone identification task on a personal computer running E-PRIME with a keyboard and AKG k141 headphones. Stimuli were presented randomly. Tones were represented as pictures of arrows. The height of an arrow within its box represented the tone’s pitch height; its direction corresponded to the tone’s pitch direction.

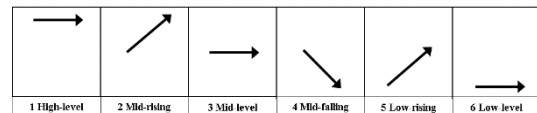


Figure 2: Visual representation of the Cantonese tones

In each trial, participants simultaneously heard a syllable and saw the arrows. Via button-press, they indicated which arrow matched the tone they heard. This abstract visual representation allowed naïve listeners to identify tones by

pitch height and direction with very little training. Each trial included a pre-response blank slide (500 ms); a response slide with the arrows and syllable (3.5 s); and a post-response blank slide (500 ms). Thus, each trial lasted 1-4.5 s. Accuracy scores (proportion correct) and reaction-time data (ms) were collected, but due to space limitations, only accuracy data are discussed. Missing responses were excluded from analysis.

Before testing, all groups, including Cantonese listeners (henceforth *Cant. listeners*), completed a tutorial on both the Cantonese tone system and the task itself. Training and practice used non-test syllables *gu*, *gi* and *ga*. Stimuli and procedure were otherwise identical to the test. Cantonese listeners also completed a 2AFC realword-identification task to establish a native-tone-identification baseline. Listeners saw a chart with a meaningful CV word (1) written in both Chinese and Roman letters with tone marks and numbers; (2) defined, in English; and (3) used in a sentence in Chinese, with English gloss. Listeners heard a tonal minimal pair and indicated which was the word in the chart.

### 3. Results

Test-condition accuracy data were submitted to two-way repeated-measures ANOVAs (rANOVAs) with L1-Group (x4) as the between-subjects factor and Tone (x6) as the within-subjects factor. Tone-identification error patterns were assessed via rANOVAs with L1-Group (x4) as the between-subjects factor and Response-Pattern (% correct vs. % each incorrect response) as the within-subjects factor.

The accuracy data show sig. main effects of Tone ( $F(4.129, 210.579)=40.627, p<0.001$ ) and L1-Group ( $F(1, 51)=18.469, p<0.001$ ) and a sig. Tone x L1-Group interaction ( $F(12.387, 210.579)=6.096, p<0.001$ ). Post-hoc Bonferroni tests indicated sig. between-group differences in mean accuracy scores for all but CLR and CLL ( $ps>0.05$ ). CLR and CLL mean scores were sig. lower than those for other tones. CLL was identified less accurately than CLR ( $ps<0.001$ ).

The effect of L1-Group on mean accuracy scores was examined via an independent-samples Kruskal-Wallis H test. L1 groups' mean accuracy score distributions were similar for CHL only: the medians of the mean accuracy scores for CHL were sig. different between groups,  $\chi^2(3)=21.301, p<0.001$ . L1 groups' mean accuracy score distributions were dissimilar for the other tone conditions, and sig. different between groups, for CMR,  $\chi^2(3)=25.943, p<0.001$ ; CML,  $\chi^2(3)=11.639, p<0.001$ ; and CMF,  $\chi^2(3)=18.304, p<0.001$ ; they were N.S. for CLR ( $\chi^2(3)=2.956, p=0.398$ ) and CLL ( $\chi^2(3)=7.048, p<0.001$ ). Pairwise comparisons using Dunn's (1964) procedure (Bonf. corr.  $\alpha=0.0083$ ) showed that Mand. listeners displayed strong Cantonese-tone identification abilities relative to other non-native listeners. Mand. and Cant. listeners outperformed Thai listeners on CHL ( $ps<0.001$ ). Mand. listeners outperformed all other groups on CMR and outperformed Yor. listeners on CML and CMF ( $ps<0.001$ ).

Table 1 shows tone confusions. Values are % of total responses. Values in bold are % of correctly-identified tones; italicized values are not sig. different from correctly-identified tone values. Asterisks indicate a main effect of Response-Pattern and a significant Response-Pattern x L1-Group interaction. For each target tone condition, rANOVAs revealed a main effect of Response-Pattern [*CHL*:  $F(2.098, 106.975)=87.051, p<0.001$ ; *CMR*:  $F(2.617, 133.465)=93.665, p<0.001$ ; *CML*:  $F(2.409, 122.877)=119.916, p<0.001$ ; *CMF*:  $F(1.573, 80.214)=92.651, p<0.001$ ; *CLR*:  $F(2.942, 150.040)=35.803, p<0.001$ ; *CLL*:  $F(2.449, 124.888)=111.719, p<0.001$ ] and a sig. Response-Pattern x L1-Group interaction [*CHL*:  $F(6.293, 106.975)=10.662, p<0.001$ ; *CMR*:  $F(7.851,$

$133.465)=8.378, p<0.001$ ; *CML*:  $F(7.228, 122.877)=4.178, p<0.001$ ; *CMF*:  $F(4.718, 80.214)=6.243, p<0.001$ ; *CLR*:  $F(8.826, 150.040)=6.815, p<0.001$ ; *CLL*:  $F(7.346, 124.888)=3.734, p<0.001$ ]. Acoustic analyses verified that the CML tokens were higher than the CLL tokens: CML mean F0 was 0.85 ST higher; CML was higher at tonal onset, midpoint, and offglide (0.87 ST, 1.06 ST, 0.52 ST, respectively).

Table 1. Cantonese tone ID confusion matrices.

Target	CANTONESE listeners' responses					
	CHL	CMR	CML	CMF	CLR	CLL
CHL*	<b>70.22</b>	16.60	6.18	4.10	1.67	1.25
CMR*	3.48	<b>42.65</b>	<i>19.48</i>	11.72	19.93	2.75
CML*	7.87	7.36	<b>51.83</b>	12.84	9.29	10.81
CMF	0.42	3.09	6.64	<b>49.49</b>	7.05	<i>33.31</i>
CLR*	3.27	<i>24.85</i>	<i>29.81</i>	<i>10.85</i>	<b>22.45</b>	8.78
CLL*	2.48	7.14	50.36	<i>14.08</i>	6.54	<b>19.40</b>
Target	THAI listeners' responses					
	CHL	CMR	CML	CMF	CLR	CLL
CHL*	<b>24.75</b>	<i>36.31</i>	<i>13.13</i>	2.19	<i>21.89</i>	<i>1.74</i>
CMR*	10.85	<b>42.16</b>	<i>15.33</i>	4.88	<i>24.12</i>	2.67
CML*	5.13	2.14	<b>73.86</b>	5.56	1.29	12.05
CMF	0.86	1.41	7.09	<b>48.63</b>	1.36	<i>40.66</i>
CLR*	6.96	<i>28.75</i>	<i>19.10</i>	7.81	<b>30.18</b>	7.21
CLL*	6.73	0.43	63.97	<i>9.00</i>	2.16	<b>17.71</b>
Target	MANDARIN listeners' responses					
	CHL	CMR	CML	CMF	CLR	CLL
CHL*	<b>75.56</b>	1.11	20.37	1.11	0.00	1.85
CMR*	1.87	<b>82.88</b>	1.13	0.00	14.12	0.00
CML*	12.85	2.27	<b>73.66</b>	0.75	1.48	9.00
CMF	0.42	0.74	3.42	<b>75.03</b>	2.68	17.72
CLR*	2.59	60.33	3.44	0.00	<b>32.90</b>	0.74
CLL*	<i>10.02</i>	3.33	63.53	4.10	1.85	<b>17.17</b>
Target	YORUBA listeners' responses					
	CHL	CMR	CML	CMF	CLR	CLL
CHL*	<b>45.18</b>	22.77	<i>16.87</i>	6.52	6.73	1.93
CMR*	3.85	<b>39.73</b>	<i>14.31</i>	13.97	26.67	1.47
CML*	14.77	16.43	<b>45.67</b>	8.76	10.25	4.13
CMF	3.78	1.42	<i>13.27</i>	<b>27.42</b>	4.68	<i>49.45</i>
CLR*	4.20	<i>21.47</i>	<i>12.80</i>	<i>23.04</i>	<b>34.63</b>	3.87
CLL*	<i>9.90</i>	<i>11.46</i>	40.00	<i>19.51</i>	<i>10.28</i>	<b>8.85</b>

## 4. Discussion

### 4.1. Cantonese tone ID in native listeners

On the realword identification post-test, the Cant. listeners hit ceiling on CHL, CMF, and CML trials (93%, 98%, 96% corr., respectively). They less-accurately identified CLL and CLR (56%, 44% corr.). CMR was identified with 71% accuracy.

On the test condition, the Cant. listeners confused CLL-CMF with each other/bidirectionally (although Response-Pattern x L1-Group interaction was N.S. for CMF) and CLR-CMR bidirectionally ( $p>0.05$ ); they confused CLL-CML unidirectionally (CML IDed correctly,  $p<0.05$ , but CLL was mis-identified as CML sig. more often than was identified correctly,  $p<0.05$ ). These results might, in part, reflect the crowded lower pitch range of the Cant. tone system and resulting acoustic overlap of CMR, CMF, CLR, and CLL, which would cause these tones to be difficult for listeners to differentiate [27]. Likewise, the post-test data suggest that CLR and CLL in particular are confusable even in realwords presented with context. Furthermore, CLL could be increasingly perceptually similar to CML and CMF, due to ongoing CLL-CML and CLL-CMF tonal mergers, while CLR-CMR is also merging [ibid.]. The Cant. listeners also appeared to have particular difficulty identifying their rising tones, confusing CLR with all but CHL (CLR-CHL  $p<0.05$ ; CLR-other tone  $p>0.05$ ) and CMR with CML ( $p>0.05$ ). Perhaps as the CLR and CMR categories merge, the new rising-tone category exemplar resembles CMR more closely than CLR along one or more acoustic dimensions. This could explain why the CLR-CMR confusion pattern was unidirectional. Thus, consistent with previous research [27], native listeners' confusions of isolated L1 tones reflected effects of acoustic similarities partly caused by ongoing tonal mergers.

### 4.2. Cantonese tone ID in non-native listeners

Consistent with predictions of the PAM-S, the non-native listeners' results reflected phonetic (dis-)similarities between L1 and novel tones, as well as (in)sensitivity to acoustic differences. Generally, the non-natives accurately identified Cant. tones similar in F0 height, direction, and/or endpoint F0 to L1 tones; they often confused tones that are markedly acoustically dissimilar from L1 tones. All non-native groups displayed a unidirectional CLL-CML confusion pattern (CML identified correctly,  $p<0.05$ ; CLL mis-identified as CML more often than identified correctly,  $p<0.05$ ). This was expected, since these tones are also confused by native listeners. The acoustic overlap (in offglide F0 and degree of F0 change) between these low-range tones could cause acoustic-perceptual confusion in naïve listeners. If it happened that CML and CLL were compared to L1 tones, they perhaps assimilated to TN (Thai), MHL (Mand.), and YML (Yor.) (CG assim.; CML the better match). (CML and TN are similar in onset F0 and F0 height/direction; CML and MHL are similar in F0 direction; CML and YML are similar in F0 height.) The Mand. and Thai groups might each have compared these two novel level tones to its single truly-level L1 tone category. As further evidence that CML and CLL were perceptually more like YML than the low-falling YLL tone, the Yor. group mis-identified CLL as a mid-range (as opposed to a falling) tone in nearly 71% of the trials.

As expected, the Thai listeners mis-identified CHL; they confused it with each of the other tones ( $p>0.05$ ). CHL might have been *uncategorizable* due to its acoustic-perceptual dissimilarity from the similarly-labeled THL in particular. The

Thai listeners also confused CMR with CML and CLR ( $p>0.05$ ). This confusion was bidirectional for CMR-CLR ( $p>0.05$ ) and unidirectional for CMR-CML (CML not mis-identified as CMR,  $p<0.05$ ). They also confused CLR-CML unidirectionally (CLR mis-identified as CML,  $p>0.05$ , but not v. versa,  $p<0.05$ ). The Thai listeners might have compared CMR, CLR, and CML to TN (CG assim. with CML being the best match). This interpretation is supported by the multiple acoustic-phonetic similarities between CML and TN (F0 height, direction, onset F0, and offglide F0) as well as the Thai listeners' strong performance on CML trials and their propensity to confuse CMR and CLR with CML and each other. Another interpretation is that they mapped CML to TN (TC assim.), and both CMR and CLR to THL, due to their F0-direction similarities (SC assim., given the bidirectional CMR-CLR confusion). Regardless, the acoustic differences between the novel target tones and familiar native tones likely contributed to the Thai listeners' perceptual confusions.

The Yor. listeners confused CMR-CLR bidirectionally ( $p>0.05$ ) and CMR-CML unidirectionally (CMR mis-identified as CML,  $p>0.05$ , but not v. versa,  $p<0.05$ ). They mis-identified CLR as CMR, CML, and CMF (all  $p>0.05$ ); also, CMR and CML were confused with CHL ( $p>0.05$ ). CMR and CLR were possibly *uncategorized* due to their acoustic-perceptual dissimilarities from any native tone categories (no isolated Yoruba tones rise steeply). CML might have been compared to YML, given its resemblance in F0 direction and mean and endpoint F0 values (TC assim.). Likewise, CHL may have been perceptually equivalent to YHL with its comparable F0 offglide value (TC assim.). The CHL tokens might simply be perceived by Yor. listeners as resembling oddly-high tokens of YHL (thus with reference to non-contrastive, *phonetic*, detail in the L1). Ongoing work attempts to disambiguate such results by asking listeners to explicitly map non-native tones to native categories.

The Mand. listeners outperformed Thai listeners on CHL and the Yor. listeners on CML (all  $p<0.001$ ). As in previous work [16, 27], they confused CLR-CMR, but unidirectionally (CMR correctly identified,  $p<0.05$ , CLR mis-identified as CMR more than correctly identified,  $p<0.05$ ). In PAM-S terms, this could be CG assim. of CLR and CMR to MR. MR resembles CLR in overall contour, but its mean F0 is lower. In contrast, MR closely resembles CMR in F0 height and direction. Since the Mand. listeners' peripheral auditory systems are experientially "tuned" to the acoustic properties of their L1 tones, the minor acoustic discrepancies between CLR and MR might have contributed to the listeners' less-accurate perception of CLR. The Mand. listeners also outperformed all other groups on CMR ( $p<0.001$ ). This is unsurprising: compared to TR, MR is more similar to CMR in F0 height, direction, slope, and offglide; Yoruba, for its part, has no comparably rising tone. (Note: One Mand. listener reported exposure to Cantonese, but exclusion of her data does not change the group averages.)

## 5. Conclusions

The current study examined identification of novel (Cantonese) tone contrasts by listeners with different tonal L1 experience. As discussed, we replicated previous reports [27] of native-Cant. CMR-CLR, CML-CLL, and CMF-CLL confusions. Cant. listeners' CLL and CLR confusions also echoed their post-test CLL and CLR realword misidentifications. These data provide additional evidence for native perceptual difficulties emerging from ongoing tonal mergers and a crowded lower acoustic tone space. Our Cant. listeners did not identify CMF as well as would be expected

from their post-test scores, but the test task, with its abstract visual tone representations and nonwords, might not have provided them enough context to consistently make overt tone-word associations in those trials. The CMF trials might therefore have involved a greater degree of lower-level acoustic processing, similar to that expected of – and demonstrated by – the non-natives. For instance, the Mand. listeners' experience processing MHL and MR, and the Thai listeners' experience with TN, might have tuned their peripheral auditory processing mechanisms in such a way as to facilitate lower-level processing of acoustically-similar novel tones (CHL and CMR resemble MHL and MR, respectively, in F0 height/direction; CML resembles TN in F0 height/direction and endpoint F0). The non-natives might also have assimilated such tones to L1 tone categories by attending to phonetic cues relevant in the L1. In contrast, some novel tones might have been sufficiently acoustically dissimilar from L1 tones as to be uncategorizable. Results suggest that non-native perception of tone is affected by experience processing acoustic-phonetic properties of L1 tone contrasts [10-11].

## 6. References

- [1] Best, C. T. (1995). A direct realist view on cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience*. Timonium, MD: York Press, 171-204.
- [2] Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 229-273.
- [3] Best, C.T., & Tyler, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.S. Bohn & M.J. Munro (Eds.), *Second language speech learning: The role of language experience in speech perception and production*, in honor of James E. Flege. Amsterdam: John Benjamins, 13-34.
- [4] Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., Trent-Brown, SA (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *JAcSocAm* 124,576-88.
- [5] Levy, E.S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American English learners of French. *J Ac Soc Am*, 125, 1138-52.
- [6] Best, C.T., McRoberts, G.W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *J Ac Soc Am*, 109, 775-794.
- [7] Bohn, O-S, & Best, CT (2012). Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *JPhon*20, 109-28.
- [8] Archibald, J. (1992). Transfer of L1 parameter settings: Some empirical evidence from Polish metrics. *Can J Ling* 37: 301-339.
- [9] Guion, S.G., Harada, T., & Clark, J.J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Biling: Lang Cog* 7:207-226.
- [10] So, C.K., & Best, C.T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Lang Sp*, 53(2), 273-293.
- [11] So, C.K. (2012). Cross-language categorization of monosyllabic foreign tone: Effects of phonological and phonetic properties of native language. In T. Stolz, N. Nau, & C. Stroh (Eds.), *Monosyllables: From phonology to typology*. Berlin, Germany: Akademie Verlag, 55-69.
- [12] So, C.K. & Best, C. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Stud Sec Lang Acq* 36:195-221.
- [13] Wu, X., Munro, J. M., & Wang, Y. (2014). Tone assimilation by Mandarin and Thai listeners with and without L2 experience. *J Phon* 46: 86-100.
- [14] Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Xu Rattanasone, N., Schwarz, I-C., & Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *App Psycholing* 36(6), 1459-1491.
- [15] Reid, A., Burnham, D., Kasisopa, B., Reilly, R., Attina, V., Xu Rattanasone, N., & Best, C. (2015). Perceptual assimilation of lexical tone: The roles of language experience and visual information. *Att Percep Psych* 77(2): 571-591.
- [16] Francis, A.L., Ciocca, V., Ma, L., & Fenn, K. (2008). Linguistic experience and the acquisition of lexical tones. *J Phon*, 36, 268-294.
- [17] Khouw, E. & Ciocca, V. (2006). Acoustic and Perceptual Study of Cantonese Tones Produced by Profoundly Hearing-Impaired Adolescents. *Ear Hear* 27(3), 243-255.
- [18] Gandour, J. (1981). Perceptual dimensions of tone: Evidence from Cantonese. *J Chin Ling*, 9: 20-36.
- [19] Zsiga, E. & Nitisaroj, R. (2007). Tone features, tone perception, and peak alignment in Thai. *Lang Sp* 50(3):343-383.
- [20] Chao, Y.R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Cambridge, MA: Harvard University Press.
- [21] Maddieson, I. (1978). Universals of tone. In *Universals of Human Language 2: Phonology* [Ed. J.H. Greenberg, et al.]. Stanford University Press, Stanford: 335-363.
- [22] Akinlabi, A. & Liberman, M. (2000). "The Tonal Phonology of Yoruba Clitics." In *Clitics in phonology, morphology and syntax*. B.Gerlach and J. Grijzenhout [Eds.] Amsterdam: Benjamins, 31-62.
- [23] Best, C.T., McRoberts, G.W., & Sithole, N.N. (1988). The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English speaking adults and infants. *J Exp Psych: HPP*, 14, 345-360.
- [24] Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A.M., Jenkins, J.J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Percep Psych*, 18, 331-340.
- [25] Best, C.T., McRoberts, G.W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *J Ac Soc Am*, 109, 775-794.
- [26] Massaro, DW, Cohen, MM, & Tseng, C (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *J Chin Ling* 13, 267-290.
- [27] Qin, Z., & Mok, P.-K. (2011). Discrimination of Cantonese tones by Mandarin, English and French speakers. *Psycholinguistic Representation of Tone*, Hong Kong.
- [28] Leather, J. (1987). F0 pattern inference in the perceptual acquisition of second language tone. In A. James & J. Leather [Eds.], *Sound Patterns in Second Language Acquisition*. Dordrecht: Foris Publications, 59-81.
- [29] Stager, J.R., & Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *J Chin Ling.*, 21, 143-163.
- [30] Saravari, C. & Imai, S. (1983). Perception of tone and short-long judgment of vowel variants of a Thai monosyllabic sound. *J Phon* 11, 231-242.
- [31] Burnham, D. & Francis, E. (1997). The role of linguistic experience in the perception of Thai tones. In Abramson A, [Ed.] *Southeast Asian Linguistic Studies in honour of Vichin Panupong*. Chulalongkorn University Press; Bangkok.
- [32] Abramson, A.S. (1978). Static and dynamic acoustic cues in distinctive tones. *Lang Sp* 21(4), 319-325.
- [33] Abramson, A.S. (1975). The tones of central Thai: Some perceptual experiments. In: J. G. Harris, & J. R. Chamberlain (Eds.), *Studies in Tai linguistics in honor of William J. Gedney*. Bangkok: Central Institute of English Language, 1-16.
- [34] Hombert, J.-M. (1976). Perception of bisyllabic nouns in Yoruba. In Hyman, L., et al. [Eds.] *Papers in African Linguistics in Honor of W.E. Welmers*, supplement 6, 109-122.
- [35] Alexander, J.A., Wong, P.C.M., and Bradlow, A.R. (2005). Lexical tone perception in musicians and non-musicians. *Proceedings of Interspeech – Eurospeech*, Lisbon, Portugal.
- [36] Blicher, D.L., Diehl, R.L., & Cohen, Leslie B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: Evidence of auditory enhancement. *J Phon* 18(1), 37-49.