10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

How sleep-mediated memory consolidation modulates the generalization across talkers: evidence from tone identification

Zhen Qin[1], *Caicai Zhang*[2]

[1] Hong Kong Polytechnic University, Hong Kong
[2] Hong Kong Polytechnic University, Hong Kong
zhen-quentin.qin@polyu.edu.hk, caicai.zhang@polyu.edu.hk

Abstract

Recent studies showed that sleep-mediated memory consolidation facilitated learners' generalization across talkers in their perception of novel stop contrasts. Lexical tone is characterized by high variability across talkers. Thus a similar effect of overnight consolidation could be found for perceptual learning of novel tonal contrasts. This study aims to examine whether overnight consolidation facilitates talker generalization in the identification of novel Cantonese level tones by Mandarin listeners.

Two groups of Mandarin listeners were perceptually trained either in the morning or in the evening using stimuli from one talker. Their post-training changes and generalization to a novel talker were then tested in three posttests over 24 hours using stimuli from the trained and untrained talkers. The results showed that the evening group showed an improved trend in identifying the level tones produced by both the trained and untrained talkers; in contrast, the morning group showed a declining trend. The finding of identification changes over time suggests that overnight consolidation might have assisted learning of tone stimuli produced by the novel talker, and eventually facilitated the formation of a more talker-independent representation of novel tone categories in long-term memory. The findings have implications for understanding the mechanism of speech learning and plasticity.

Index Terms: overnight consolidation, Cantonese tones, Mandarin, perceptual learning

1. Introduction

Converging evidence indicates that sleep supports various aspects of language learning by facilitating the memory consolidation of newly learned knowledge [1], [2]. Sleep-mediated memory consolidation (i.e., overnight consolidation) plays an important role in novel word learning [2]–[6]. On the other hand, a growing literature showed that overnight consolidation also facilitates listeners' perceptual learning of speech sound categories in an auditory domain [7]–[11].

In recent studies, Earle and her colleagues examined how learned sound categories were encoded into long-term memory following a session of laboratory training by focusing on the effect of overnight consolidation in the identification and discrimination of novel phonetic categories, that is, a Hindi dental and retroflex stop contrast [12]–[16]. Specifically, Earle and Myers showed that overnight consolidation promoted English listeners' generalization across talkers in their identification, but not in the discrimination, of the novel Hindi stop contrast [13]. The results of identification tests showed that while English listeners trained in the evening improved significantly in identifying the stop stimuli produced by the untrained talker, those trained in the morning did not show such a pattern. In contrast, the two training groups did not show perceptual changes in identifying the stop stimuli produced by the trained talker.

The overnight consolidation effect in talker generalization (i.e., talker-independent abstraction) found for the perceptual learning of non-native consonants, albeit informative, raises the question of whether a similar effect would be found in the perceptual learning of lexical tones, which are highly variable and dynamic within and across talkers [17], [18]. Tone languages use pitch variations to distinguish lexical meanings. For instance, segmentally identical words that contain different lexical tones in Cantonese (/si 55/ 'silk' (Tone 1 (T1)), /si 25/ 'history' (T2), /si 33/ 'to try' (T3), /si 21/ 'time' (T4), /si 23/ 'city' (T5), and /si 22/ 'matter' (T6)) differ in meaning [19].

Cross-linguistic studies found that listeners with different tone language backgrounds attend to different pitch cues under the influence of their L1 tonal inventory [20]–[22]. Cantonese, in particular, has contour tone contrasts as well level tone contrasts. In contrast, Mandarin has one level tone and three contour tones, and thus contrast in pitch contour differences. Some Cantonese level-level tone pairs, for instance, T3 (mid-level) – T6 (low-level), are distinguished in subtle differences in pitch height, and thus pose a great perceptual difficulty for Mandarin listeners [23], [24]. Since Mandarin listeners have been reported to have a great difficulty distinguishing (e.g., Cantonese) level-level tonal contrasts [25]–[27], this study examines whether, and if so how, a talker generalization in perceptual learning of novel tones, that is, three Cantonese level tones (T1 /55/, T3 /33/, and T6 /22/) learned by Mandarin listeners, is modulated by the effect of overnight consolidation.

Given previous findings on the learning of novel segmental contrasts, it is hypothesized that a similar effect of overnight consolidation will be found for the perceptual learning of Cantonese level tones by Mandarin listeners. Specifically, those listeners who are trained in the evening are expected to perform better than those who are trained in the morning in perceiving the level tones produced by a new talker and probably also the trained talker, given a potentially greater difficulty of learning the target (three-way) tonal contrasts relative to (two-way) segmental contrast. Lastly, it is expected that the overnight consolidation effect would be across-the-board for the perceptual learning of all three level tones.

10.21437/SpeechProsody.2020-94

2. Methods

2.1. Participants

Thirty-two students (18 female, 14 male) were recruited from the Hong Kong Polytechnic University. They were all Mandarin-speaking participants between the ages of 18 and 30. The participants had a minimal exposure to Cantonese, that is, their length of residence in Hong Kong was less than ten months and Cantonese classroom-learning for less than one month prior to the experiment. They speak Mandarin as their mother tongue and not any Southern Chinese dialect including Shanghainese, Hakka, and Southern Min. In addition, they have not received professional musical training. None of them reported a history of hearing impairment and neurological disorder.

Table 1: *Demographic characteristics of the morning and evening training groups.*

		Morning	Evening
No. of participants		16 (9 F, 7 M)	16 (9 F, 7 M)
Age (year)		24.9 (3.7)	23.3 (3.5)
Discrimination (%)	Pretest	85.2 (3.6)	84.5 (3.6)

The participants were equally and randomly assigned into one of the two groups who were trained either in the morning (8-10 am) or evening (8-10 pm). Demographic characteristics of the morning and evening groups are summarized in Table 1. To ensure that the morning and evening groups had comparable levels of tone discrimination sensitivity before going through training, an AX discrimination pretest a baseline assessment was conducted. The accuracy of the two groups did not significantly differ from each other [t (30) = –.46, p = .65].

2.2. Stimuli

The stimuli were three Cantonese level tones, /55/ T1 (a high-level tone), /33/ T3 (a mid-level tone), and /22/ T6 (a low-level tone) carried by ten base syllables (/jan/, /ji/, /jau/, /jiu/, /fan/, /fu/, /ngaa/, /si/, /se/ and /wai/). All thirty words are meaningful in Cantonese. Two female speakers of Hong Kong Cantonese recorded three repetitions of each target word. Each monosyllabic target word was recorded in a carrier phrase "lei1 go3 hai6 [target word]" (this is [target word]). Recordings were conducted in a sound-proof room using a microphone linked to a digital recorder. Each token was segmented out of the carrier phrase. To increase the variability of tone stimuli, two tokens for each target word were chosen by the investigators based on its intelligibility and sound quality. The stimuli were normalized in duration to 500 ms (a value similar to the duration of naturally produced stimuli), and their mean acoustic intensity was scaled to 70 dB using Praat.

As shown in Figure 1, the untrained talker, with a higher average pitch and a narrower pitch range, differed phonetically from the trained talker in their pitch distribution, and thus was used as a generalization talker in the posttest session alone.

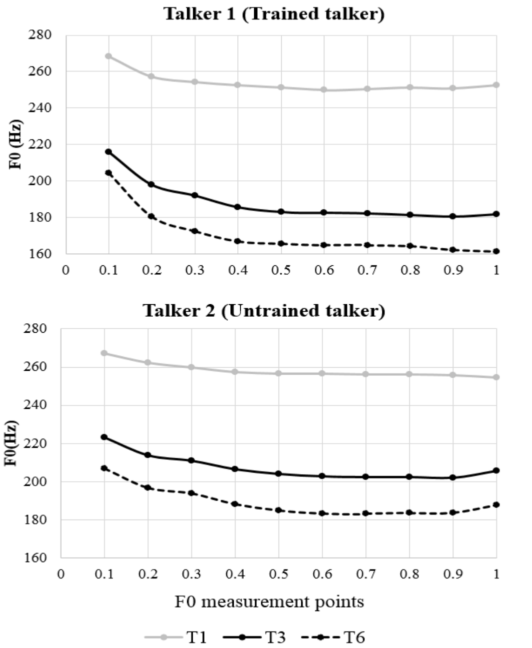

Figure 1: *Tonal contours of the three Cantonese level tones produced by the trained talker (top panel) whose stimuli were used in both the training and posttests as well as the untrained talker (bottom panel) whose stimuli were used in the posttests alone.*

2.3. Procedures

Participants completed a pretest, a training, and a posttest in three sessions over two consecutive days. In the training session, a forced-choice identification (ID) task of three Cantonese level-tone categories was conducted on the participants in a separate day. In case that the training stimuli might be perceptually challenging, only monosyllables produced by the first female speaker, who had a wider range of pitch than the other speaker, were used in the training session. A total of 300 tokens (1 speaker * 3 tones * 10 syllables * 2 tokens * 5 repetitions) were presented in auditorily to the participants in five blocks with 60 tokens in each block. During the training, the participants were instructed to identify each tone (T1-High, T3-Mid and T6-Low) after hearing the auditory stimuli by pressing three buttons (1, 3, and 6) in a self-paced fashion. Written feedback ("Correct" in green or "Incorrect. The correct answer is…" in red) was given immediately after every trial. The participants were instructed to learn to categorize three tones based on feedback, and achieve the best performance as they can in the training session. The training session took approximately 30 min.

In the posttest sessions, each posttest was conducted at three time points (immediately after training, with a 12-hour delay, and with a 24-hour delay) to test how the learned Cantonese level tones were retained and how the representation of learned tones changed over time after perceptual training. Monosyllables produced by both the trained (Talker 1 in Figure 1) and untrained (Talker 2 in Figure 1) talkers were used in the posttest sessions. In the ID posttests, the participants were instructed to identify each tone (T1-High, T3-Mid and T6-Low) by pressing three buttons (1, 3, and 6). However, no feedback was given after every trial. The ID posttest consisted of 120 trials (2 talkers * 3 tones * 10 syllables * 2 tokens), which were randomly presented to participants in one block. The ID posttest took approximately 10 min.

2.4. Data Analysis

A logit mixed-effects model with Group (2 levels: morning vs. evening; morning as Baseline), Time (3 levels: ID posttest1, ID posttest2, ID posttest3; ID posttest1 as Baseline), Talker (2 levels: trained vs. untrained; trained talker as Baseline), and Tone (3 levels: T1, T3, and T6; T1 as Baseline) as fixed effects, with participants and items as random intercepts and Time as a random slope on the participant random intercept (the models with more random slopes did not converge), was performed on the participants' accuracy data (1 = correct, 0 = incorrect) of their ID posttests.

Mixed-effects models with a maximal random effects structure used in the current study have greater statistical power than the analysis of variance, and are more appropriate for our binomial data of identification accuracy [28]. The models were fitted in R, using the lmer() function from the lme4 package [29]. A back fitting function from the package *LMERConvenienceFunctions* in R was used to identify the best model that accounted for significantly more of the variance than simpler models, as determined by log-likelihood ratio tests [30]; only the results of the model with the best fit are presented, with *p* values calculated using the *lmerTest* package in R [31].

3. Results

Figure 2 presents the morning and evening training groups' mean proportions of correct response (i.e., accuracy) for stimuli produced by the trained and untrained talkers in the ID posttests over the 24-hour period (ID posttest1, ID posttest2, and ID posttest3).

Recall that a logit mixed-effects model was performed on participants' accuracy to examine the effects of Group (morning vs. evening), Time (ID posttest1, ID posttest2, ID posttest3), Talker (trained vs. untrained), and Tone (T1, T3, and T6), and their interactions. The baseline was the morning groups' performance on trained stimuli of T1 in the ID posttest1. The model with the best fit included the simple effects of Group, Time, Tone, Talker, as well as the interaction between Group and Time. The estimate, standard error, *z* value, and *p* value associated with the fixed effects are presented in Table 2.

The model results summarized in Table 2 yielded a significant two-way interaction between Group and Time, indicating that the two groups differed in their ID performance changes over the 24-hour experiment period. Crucially, the three-way interaction among Group, Time and Talker did not improve the model, indicating a lack of evidence that the ID performance changes over time of the two groups differed between stimuli produced by the trained and untrained talkers. The lack of interaction between Tone and other factors indicates that ID performance changes (e.g., for the morning/evening groups and for trained/untrained stimuli) remained similar for all the level tones.

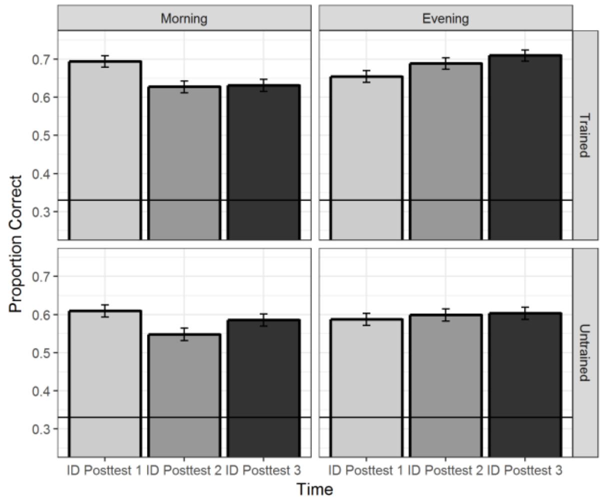

Figure 2: *The morning and evening training groups' mean proportions of correct response for stimuli produced by the trained and untrained talkers over the 24-hour period in the ID posttests. The error bars represent one standard error of the mean; the horizontal line represents chance performance.*

To understand the nature of the significant two-way interaction between Group and Time, subsequent logit mixed-effects models were therefore performed on the participants' accuracy across three ID posttests separately for each training group.

Table 2: *Best logit mixed-effects model on proportions of correct response of the participants from the morning and evening groups in the three ID posttests.*

Effect	**Est.**	**SE**	***z***	***p***
(intercept)	1.6	0.14	10.9	< .001
Group	–0.2	0.19	–0.9	.34
Time (ID Posttest2)	–0.3	0.09	–3.8	< .001
Time (ID Posttest3)	–0.2	0.08	–2.7	< .01
Tone	–0.2	0.01	–19.4	< .001
Talker	–0.4	0.04	–9.2	< .001
Group × Time (ID Posttest2)	0.4	0.13	3.6	< .001
Group × Time (ID Posttest3)	0.4	0.12	3.4	< .001

For the morning group, the participants' performance in either the ID posttest2 (Estimate = –0.320, Std. Error = 0.087, $z = –3.68$, $p < .001$) or the ID posttest3 (Estimate = –0.220, Std. Error = 0.072, $z = –2.99$, $p < .01$) was significantly lower than their performance in the ID posttest1. However, the participants' performance between the ID posttest2 and the ID posttest3 did not significantly differ (Estimate = 0.106, Std. Error = 0.070, $z = 1.51$, $p = .13$). The results of the morning group suggest that participants without an intervening night's sleep between Session 1 (i.e., training session) and Session 2 (i.e., posttest2) showed a significant declining ID performance in the later posttests compared with their ID performance in the initial posttest.

For the evening group, the participants' performance in the ID posttest2 did not differ from their performance in the ID posttest1 (Estimate = 0.109, Std. Error = 0.081, z = 1.33, p = .18). The participants' performance difference between the ID posttest3 and the ID posttest1 had a marginal significance (Estimate = 0.153, Std. Error = 0.082, z = 1.88, p = .06). Again, the participants' performance in the ID posttest2 and the ID posttest3 did not significantly differ (Estimate = 0.043, Std. Error = 0.071, z = 0.61, p = .54). The results of the evening group suggest that participants with an intervening night's sleep between Session 1 (i.e., training session) and later sessions showed a numerical trend of improved ID performance in posttest2, and a marginally significant improvement in posttest3, compared with their ID performance in the initial posttest. However, there is no evidence that their ID performance improved during the 12 waking hours between posttest2 and posttest3. Importantly, the different perceptual patterns over time between the morning and evening training groups were found for the stimuli produced by both the trained and untrained talkers.

4. Discussion and Conclusion

The results of identification tasks showed that the Mandarin listeners who were trained in the morning, without an intervening night's sleep between training and Posttest2, showed a declining performance of identification accuracy; in contrast, the Mandarin listeners who were trained in the evening, with an intervening night's sleep between training and later posttests, showed a numerical trend of improved performance in identification accuracy. Crucially, the pattern was found for stimuli produced by both the trained and untrained talker across all three level tones.

First and foremost, the findings suggest that the sleep intervening training and posttests facilitated talker generalization in perceptual learning of novel tones. That is, the difference of post-training identification changes between the evening and morning trainees was not only found for stimuli produced by the trained talker, but also for stimuli produced by the novel untrained talker. This finding implies that overnight consolidation might have assisted the evening trainees' generalization to stimuli produced by the untrained talker. This finding of talker generalization seems to support an overnight consolidation process from acoustic-phonetic (i.e., pitch) features at an episodic level to a formation of a more abstract representation of novel tone categories in long-term (e.g., declarative) memory [1], [2]. This possibility is in line with the previous findings that overnight consolidation promoted listeners' speech sound abstraction in perceptual learning of non-native segments [13], [32].

More specifically, compared with the morning trainees, the sleep protected the evening trainees' identification of novel tonal contrasts against subsequent decay, and further induced a trend of improved identification performance. The finding suggests that overnight consolidation might have exerted either a protective or restorative effect on the evening trainees' post-training performance of tone identification. The protective effect found in perceptual learning of novel tones is consistent with previous findings in phonetic learning [7], [8]. In addition, the trend of improved identification performance over time found for the evening trainees is also (partially) in line with previous research on perceptual learning of segments which showed an improved identification performance after sleep [13].

Note that although there was a trend of improved identification performance over the 24-hour period for the evening trainees, it was not statistically significant. The discrepancy between this finding and that of previous studies could be accounted by the difference of target stimuli [13]. One possible explanation is that the variable and dynamic nature of lexical tones may have made them more difficult to learn than segments (e.g., stops) [33], [34]. As a result, different from previous studies (on perceptual learning of dental and retroflex stops) which found that the morning trainees did not change their identification performance over time, the morning trainees in the current study exhibited declining post-training identification performance over time. In a similar vein, this account may explain why the evening trainees exhibited a non-significant trend of improvement in their post-training identification performance, while managing to maintain the training-induced tone learning after sleep. Given the potential difficulty of tone learning, future studies may consider using a longer training session (e.g., 1 hour) to test whether improved post-training tone identification will be found at a group level an overnight consolidation.

Different from the previous study [13] which found the different perceptual changes of the two training groups in identifying stimuli from the untrained talker alone, we found divergent performance changes of the two groups in identifying stimuli from both the trained and untrained talkers. This discrepancy can also be attributed to the different nature of segments and lexical tones. As mentioned earlier, since tones might be more perpetually difficult to learn/abstract than segments. Compared with the morning groups who showed a declining identification performance for stimuli produced by the trained talker, which were already perceptually challenging, the intervening sleep between training and later posttests may have helped the evening trainees, at least, maintain their learning of the stimuli from the trained talker.

To conclude, the present study is one of the first studies to examine the effect of overnight consolidation in perceptual learning of lexical tones. The present findings suggest that overnight consolidation might facilitate the talker-independent abstraction of Cantonese level tones by promoting the tone learning from the trained talker to the untrained talker. Sleep protected evening trainees' identification from subsequent decay, and yielded identification performance changes, over the 24-hour experiment period. These findings raise several questions, for instance, the nature of target stimuli, for further research [8]. It would be intriguing to investigate the potential effect of overnight consolidation in perceptual learning of different types of tone pairs (e.g., contour-contour tone pairs vs. level-level tone pairs) and other prosodic categories (e.g., lexical stress) to further shed light on the effect of overnight consolidation in perceptual learning of suprasegmental domain as well as different nature of segments and prosodic categories in this process.

5. Acknowledgements

This work was supported in part by the *Language Learning* Early Career Research Grant awarded to the first author, and the Departmental General Research Funds (International collaboration) awarded to the second author, as well as by HKRGC-GRF 15601718. We would like to thank Prof. William S-Y. Wang and other team members for their valuable comments on this research.

6. References

[1] J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly, "Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory," *Psychol. Rev.*, vol. 102, no. 3, pp. 419–457, 1995.

[2] M. Davis, A. M. Di Betta, M. Macdonald, and G. Gaskell, "Learning and consolodidation of novel spoken words: Behavioral and neural evidence," *Cereb. Cortex*, pp. 803–820, 2008.

[3] C. Breitenstein *et al.*, "Hippocampus activity differentiates good from poor learners of a novel lexicon," *Neuroimage*, vol. 25, no. 3, pp. 958–968, 2005.

[4] V. Havas, J. S. H. Taylor, L. Vaquero, R. de Diego-Balaguer, A. Rodríguez-Fornells, and M. H. Davis, "Semantic and phonological schema influence spoken word learning and overnight consolidation," *Q. J. Exp. Psychol.*, vol. 71, no. 6, pp. 1469–1481, 2018.

[5] M. H. Davis, A. M. Di Betta, M. J. E. Macdonald, and M. G. Gaskell, "Learning and consolidation of novel spoken words," *J. Cogn. Neurosci.*, vol. 21, no. 4, pp. 803–820, 2009.

[6] N. Landi *et al.*, "Neural representations for newly learned words are modulated by overnight consolidation, reading skill, and age," *Neuropsychologia*, vol. 111, pp. 133–144, 2018.

[7] K. M. Fenn, H. C. Nusbaum, and D. Margoliash, "Consolidation during sleep of perceptual learning of spoken language.," *Nature*, vol. 425, no. 6958, pp. 614–6, 2003.

[8] K. M. Fenn, D. Margoliash, and H. C. Nusbaum, "Sleep restores loss of generalized but not rote learning of synthetic speech," *Cognition*, vol. 128, no. 3, pp. 280–286, 2013.

[9] T. P. Brawn, H. C. Nusbaum, and D. Margoliash, "Sleep Consolidation of Interfering Auditory Memories in Starlings," *Psychol. Sci.*, vol. 24, no. 4, pp. 439–447, 2013.

[10] F. Eisner and J. M. McQueen, "Perceptual learning in speech: Stability over time," *J. Acoust. Soc. Am.*, vol. 119, no. 4, pp. 1950–1953, 2006.

[11] F. S. Earle and E. B. Myers, "Building phonetic categories: An argument for the role of sleep," *Frontiers in Psychology*, vol. 5, no. OCT. pp. 1–12, 2014.

[12] F. S. Earle, N. Landi, and E. B. Myers, "Adults with Specific Language Impairment fail to consolidate speech sounds during sleep," *Neurosci. Lett.*, 2018.

[13] F. S. Earle and E. B. Myers, "Overnight consolidation promotes generalization across talkers in the identification of nonnative speech sounds," *J. Acoust. Soc. Am.*, vol. 137, no. 1, pp. EL91–EL97, Jan. 2015.

[14] F. S. Earle, N. Landi, and E. B. Myers, "Sleep duration predicts behavioral and neural differences in adult speech sound learning," *Neurosci. Lett.*, 2017.

[15] F. S. Earle and D. T. Arthur, "Native phonological processing abilities predict post-consolidation nonnative contrast learning in adults," *J. Acoust. Soc. Am.*, vol. 142, no. 6, pp. EL525–EL531, 2017.

[16] F. S. Earle and E. B. Myers, "Sleep and native language interference affect non-native speech sound learning," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 41, no. 6, pp. 1680–1695, Dec. 2015.

[17] C. Zhang and S. Chen, "Toward an integrative model of talker normalization," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 42, no. 8, pp. 1252–1268, 2016.

[18] Z. Qin, A. Tremblay, and J. Zhang, "Influence of within-category tonal information in the recognition of Mandarin-Chinese words by native and non-native listeners: An eye-tracking study," *J. Phon.*, vol. 73, 2019.

[19] M. Yip, "Tone: Phonology," in *Encyclopedia of Language & Linguistics*, 2006.

[20] J. T. Gandour, "Tone perception in far eastern-languages," *J. Phon.*, vol. 11, no. 2, pp. 149–175, 1983.

[21] I. L. Chan and C. B. Chang, "Perception of nonnative tonal contrasts by Mandarin-English and English-Mandarin sequential bilinguals," *J. Acoust. Soc. Am.*, vol. 146, no. 2, pp. 956–972, 2019.

[22] B. Li, J. Shao, and M. Bao, "Effects of Phonetic Similarity in the Identification of Mandarin Tones," *J. Psycholinguist. Res.*, vol. 46, no. 1, pp. 107–124, 2017.

[23] P. P. K. Mok, D. Zuo, and P. W. Y. Wong, "Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese," *Lang. Var. Change*, vol. 25, no. 3, pp. 341–370, 2013.

[24] K. Y. S. Lee, K. T. Y. Chan, J. H. S. Lam, C. A. Van Hasselt, and M. C. F. Tong, "Lexical tone perception in native speakers of Cantonese," *Int. J. Speech. Lang. Pathol.*, vol. 17, no. 1, pp. 53–62, 2015.

[25] Z. Qin and A. Jongman, "Does Second Language Experience Modulate Perception of Tones in a Third Language?," *Lang. Speech*, vol. 59, no. 3, pp. 318–338, Sep. 2016.

[26] Y. S. Chang, Y. Yao, and B. H. Huang, "Effects of linguistic experience on the perception of high-variability non-native tones," *J. Acoust. Soc. Am.*, vol. 141, no. 2, pp. EL120–EL126, 2017.

[27] A. L. Francis, V. Ciocca, L. Ma, and K. Fenn, "Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers," *J. Phon.*, vol. 36, no. 2, pp. 268–294, 2008.

[28] J. A. Linck and I. Cunnings, "The Utility and Application of Mixed-Effects Models in Second Language Research," *Lang. Learn.*, 2015.

[29] R. H. Baayen, "Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Harald Baayen (2008)," *Socioling. Stud.*, 2008.

[30] A. Tremblay and J. Ransijn, "LMERConvenienceFunctions: A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions. R package version 2.1.," *Compr. R Arch. Netw.*, 2015.

[31] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models. R package version 3.0-1," *J. Stat. Softw.*, 2018.

[32] X. Xie, F. S. Earle, and E. B. Myers, "Sleep facilitates generalisation of accent adaptation to a new talker," *Lang. Cogn. Neurosci.*, 2018.

[33] L. Singh, H. H. Goh, and T. D. Wewalaarachchi, "Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation," *Cognition*, vol. 142, pp. 1–11, 2015.

[34] T. D. Wewalaarachchi, L. H. Wong, and L. Singh, "Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English–Mandarin toddlers," *J. Exp. Child Psychol.*, vol. 159, pp. 16–33, 2017.