Perceptual Evaluation of Mandarin Tone Sandhi Production by Cantonese Speakers before and after Perceptual Training

Bei Li
Department of Chinese and Bilingual Studies,
The Hong Kong Polytechnic University
Hung Hom, Hong Kong
beili@polyu.edu.hk

Yike Yang
Department of Chinese and Bilingual Studies,
The Hong Kong Polytechnic University
Hung Hom, Hong Kong
yi-ke.yang@connect.polyu.hk

Si Chen
Department of Chinese and Bilingual Studies,
The Hong Kong Polytechnic University
Hung Hom, Hong Kong
sarah.chen@polyu.edu.hk

Abstract

This study is the first to examine the effect of perceptual training on Mandarin tone sandhi production by Cantonese speakers. Auditory and visual inputs of tone sandhi contrasts were included in a short-term laboratory training, and the training was followed by an identification test. Twenty-four native speakers of Cantonese participated in the study, which comprised the training session and a pre- and post-training recording session. There were 192 target stimuli of real words and wug words and 192 filler words in each recording session. Two native Mandarin-speaking linguists perceptually evaluated a total of 23040 syllables on a 101-point scale. The results show that Cantonese speakers may be able to improve their lexical word production in the context of T3+T1/T2/T4 by perceptual training or high familiarity of stimuli, whereas the application of Mandarin sandhi Tone 3 slightly improves within a short time. Besides, the participants consistently apply Mandarin half-third tone sandhi rule with a lexical mechanism rather than a computational mechanism, whereas the preference to lexical mechanism is not found in the application of Mandarin third tone sandhi rule.

1 Introduction

It has been shown that Cantonese speakers acoustically differed from Beijing Mandarin speakers in their application of two Mandarin tone sandhi rules (Chen et al., 2017). Specifically, Cantonese speakers preferred to neutralize the sandhi Tone 3 more with Tone 2 in isolation which indicated a more categorical application of the third tone sandhi rule. Also, they tended to raise f0 value at the end of sandhi Tone 3 which should be modified by half Tone 3 without the raising part.

Mandarin and Cantonese are both tonal languages, which discriminate lexical meanings by different tones. Mandarin has four lexical tones (Chao, 1948):

Tone1 (T1) high-level (55) "ma" (mother)
Tone2 (T2) high-rising (35) "ma" (hemp)
Tone3 (T3) low-dipping (213) "ma" (horse)
Tone4 (T4) high-falling (51) "ma" (to scold)

There are two rules of tone sandhi (Zhang & Lai, 2010) in Mandarin:

- Third tone sandhi rule:

T3 (213) → T2 (35)/__ T3 (213)
lau 213 - pan 213 → lau 35 - pan 213 "boss"

- Half-third sandhi rule:

T3 (213) → 21/__ {T1(55), T2(35), T4(51)}
lau 213 - ʂi55 → lau 21 - ʂi55 "teacher"
lau 213 - liəŋ 35→ lau 21 - liəŋ 35 "old age"
lau 213 - xua 51→ lau 21 - xua 51 "old saying"

Six contrastive tones are involved in Cantonese (Bauer & Benedict 1997): high-level (55), high-rising (25), mid-level (33), mid-low falling (21), mid-low rising (23), and mid-low level (22). Taking the monosyllable /ji/ with six tones as examples:

T1 high-level (55) "ji" (clothes)
T2 high-rising (25) "ji" (chair)
T3 mid-level (33) "ji" (meaning)
T4 mid-low falling (21) "ji" (son)
T5 mid-low rising (23) "ji" (ear)
T6 mid-low level (22) "ji" (two)

Studies have overwhelmingly proved that adults can acquire non-native tonal contrasts through either long-term or short-term perceptual training (Wayland & Li, 2007; Cooper & Wang, 2013; Antoniou & Chin, 2018). Wayland and Guion (2004) revealed that native Chinese speakers showed better performance in perceiving the unfamiliar Thai tones than non-tonal English speakers. Wang (2003)'s study proved that long-term perceptual training on Mandarin isolated tones improved the accuracy of American English speakers' tone production. For studies on training and evaluation of tone production, the majority of them only concerns the production of isolated tones and often excludes tone sandhi rules even when examining disyllabic context (Wang et al., 1999, 2003). Similar to perception and production, the evaluation of tones is also limited in isolated tones in previous studies (Wayland, 1997; Wang et al., 2003). Few studies have been complemented by evaluating dynamic tones with sandhi.

As a pilot study to compare the production before and after perceptual training, the current study utilized perceptual evaluation to rate the transfer of perceptual acquisition to production. It examined speech production by L2 learners of Mandarin who speak a tonal Cantonese as the first language (L1). Cantonese speakers of TG met the criterion that trainees should obtain an accuracy of 60 percentage or above in identification test to make sure that perceptual training was effectively conducted.

In current study, L1 Cantonese speakers were invited to participate in pre- and post-training recording sessions. The perceptual evaluation was to judge the degree of application of tone sandhi rules, as well as the naturalness of speech production. The goal of the current study is to address the following two questions:
1) Does perceptual training improve tone sandhi production by Cantonese speakers?
2) Does perceptual training improve two tone sandhi rules equally? If not, what are the differences?

2 Method

Two native Mandarin-speaking linguists were asked to rate the tone sandhi application degree of each tone in the disyllabic tonal combination on a 101-point scale. Cantonese speakers of training group (TG) and control group (CG) were presented by two Mandarin monosyllabic words with audio and visual stimuli by E-prime. The audio stimuli were recorded by one female native Mandarin speaker and were normalized the intensity and duration. Participants were presented with characters or phonetic symbols on the screen when hearing and pronouncing the words. Then they were asked to produce the two monosyllabic words as one disyllabic word in both pre- and post-tests. TG received a short-term laboratory perceptual training on Mandarin tone sandhi rules after the pre-test.

2.1 Participants

A total of 24 native speakers of Cantonese were recruited in the study, with 12 participants (eight females, four males; mean age ± SD: 23.00 ± 1.58 years) in the training group (TG) and 12 participants (seven females, five males; mean age ± SD: 23.10 ± 2.72 years) in the control group (CG). The trainees started their formal Mandarin courses to learn Mandarin at an average age of 18.33 years (SD=2.49) and have been learned for an average of 2.42 years (SD=1.23) as reported. For the controls, the average age of starting to learn Mandarin was 17.83 years (SD=2.11), with an average of 2.5 years (SD=0.76) of study length. The length of residence in Mandarin-speaking Mainland China is 26 days on average for the trainees, and 15.5 days for the controls. None of the participants reported a history of hearing or speaking problems. Twelve Mandarin native speakers born in Beijing also participated in the recording session.

All participants were recruited from the Hong Kong Polytechnic University and were paid for their participation. They were recorded with a GMH C 8.100 D headset at the speech lab of the Hong Kong Polytechnic University.

Two adult native Mandarin-speaking trained linguists (one female, one male; 27 and 29 years old respectively) conducted the perceptual evaluation. They have no reported hearing or speech impairments.

2.2 Stimuli

Stimuli for Production Experiment: Following Zhang and Lai (2010), we designed six types of stimuli including both real words (Type 1) and wug words (Type 2 to Type 6):

1. Real disyllabic words (AO-AO, where AO stands for actual occurring morphemes);
2. Non-occurring sequences consisting of real morphemes (*AO-AO);
3. Sequences of a real morpheme and a syllable of an accidental gap, which has a legal syllable, but bears a tone not allowed on that syllable in Mandarin (AO-AG, where AG stands for accidental gap);
4. Sequences of a syllable of an accidental gap and a real morpheme (AG-AO);
5. Sequences of two syllables of accidental gaps (AG-AG);
6. Sequences of two pseudo words, where the combination of vowels and consonants do not exist in Mandarin, but the vowel and consonant components of each syllable do exist.

In total, there were 192 target stimuli with all possible tonal combinations, as well as 192 filler words (both real and wug words) to disguise the purpose of experiment. All monosyllabic stimuli used to elicit speech production were read by a female native Mandarin speaker born in Beijing, and T3 was pronounced as a full tone with a falling-rising f0 contour.

The same sets of stimuli were used in both pre- and post-training recording sessions. Totally, the speech samples produced by each group (TG, CG and Beijing speakers) consisted of 4608 disyllabic words ((192 target disyllabic words +192 filler disyllabic words) *12 subjects), including 2304 samples of target disyllabic words for each test

Stimuli for Perceptual Training: Six real words were used for the perception training and follow-up identification test. Three of them applied Mandarin third tone sandhi rule (T3+T3: /məi/ /xau/ "美好", /ʂʷəi/ /kʷoo/ "水果" and /ʈʂʷæn/ /lʲæn/ "转脸"), and the other three applied the half-third sandhi rule (T3+T1: /ʂʷəi/ /tɕəŋ/ "水晶"; T3+T2: /ʈʂʷæn/ /ii/ "转移"; T3+T4: /məi/ / lʲii/ "美丽"). A total of 72 tokens of the six words produced by five native Mandarin female speakers was adopted in the training session. Meanwhile, there were also 72 synthesized tokens without tone sandhi as their f0 contour of the sandhi syllable was replaced by their own T3 production in isolation with normalized intensity and duration.

2.3 Procedure

Pre and Post Production Experiments: In the practice session, we first asked the participants to familiarize themselves with the materials and tested them by dictation to ensure their basic knowledge of Mandarin disyllabic pronunciation and phonetic symbols (pinyin).

Similar to previous Mandarin sandhi studies such as Zhang (2010), we used a nonce-probe paradigm ('wug' test) (Berko, 1958) to complement the production experiments. The participants heard two monosyllables together with their characters (if available) and phonetic symbols presented in E-prime, and were instructed to produce them as one disyllabic word. At the same time, they were recorded by Praat on another PC and were allowed to correct themselves. Twenty-four blocks (six types of words * four tonal combinations) of stimuli were counterbalanced across participants. The procedures were the same for pre- and post- production experiments.

Perceptual Training and Identification Test: Participants of TG attended the short-term laboratory training on Mandarin sandhi rules between pre- and post- recording sessions. The perceptual training session contained two parts. The first part was an auditory and visual explanation of tone sandhi rules, as well as an auditory-only practice. Specifically, odd trials were naturally produced sandhi stimuli and even trials were the synthesized stimuli without tone sandhi. Both scenarios were accompanied with visual f0 contours as below Figure 1 (without tone sandhi) and Figure 2 (with tone sandhi) (Chen et. al, 2017) displayed in the explanation part. Trainees were instructed to distinguish the differences of two sandhi rules, and press "A" when hearing the words with tone sandhi and "L" for words without tone sandhi. The following practice session comprised of 144 tokens with the same procedure but without pictures of f0 contours as visual input.

The other part was the identification test after training. Trainees were asked to identify the random trials and press “A” and “L” for correct and incorrect respectively. Feedback with accuracy and reaction time was immediately presented after the press. Twelve trainees of the current study obtained an accuracy of 60% or above in the test.

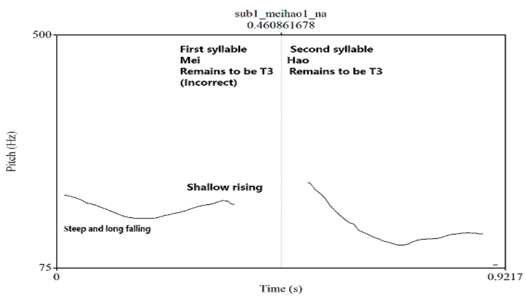

Figure1. F0 contour of “/məi3/ /xau3/”without tone sandhi

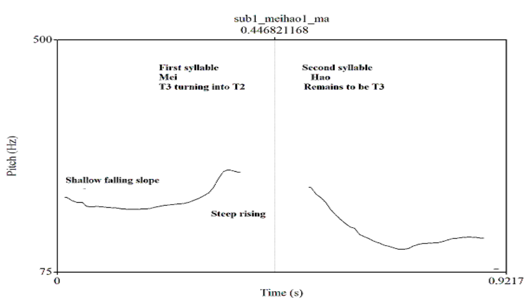

Figure2. F0 contour of “/məi3/ /xau3/”with tone sandhi

Rating Tone Sandhi Productions: All speech samples produced by all participants were first manually segmented and normalized. After anonymization, the raters were instructed to evaluate the realization of tone sandhi and avoid the effects of consonants and vowels. In other words, the first syllable of each disyllabic word was rated on a 101-point scale, which evaluated the degree of application of tone sandhi as well as the naturalness of tone production. Specifically, 100 point referred to full application of tone sandhi rules in terms of native listeners’ intuition which was close to native-like level, whereas 0 point indicated non-application of tone sandhi rules which was pronounced with a heavy non-native accent.

2.4 Data Analysis

In total, there were 23040 rated samples (2304 monosyllabic sandhi syllables * 5 times (two tests for TG and CG, one for Beijng speakers) * two raters) which included 5760 samples applying the third tone sandhi rule and 17280 samples applying the half-third sandhi rule. Results of Spearman's rank correlation rho were 0.7656 ($p < .001$) in pre-test and 0.5056 ($p < .001$) in post-test for two raters. Paired samples t-test was adopted to examine the difference between pre- and post-tests. Following sections present the results based on three factors of present study: two sandhi rules (third tone sandhi rule vs. half-third tone sandhi rule), familiarity of words (trained words vs. new words) and lexical meaning (real words vs. wug words).

3 Results

Generally, paired samples t-test results showed that both TG and CG significantly improved their tone sandhi production from pre-test to post-test as presented in Figure 3. The mean rating score of TG was 69.16 for pre-test and 72.77 for post-test ($p < .001$). CG also increased from 63.38 to 67.49 ($p < .001$). However, the performance of Cantonese speakers was obviously inferior to native speakers (M=91.16) even after the perceptual training.

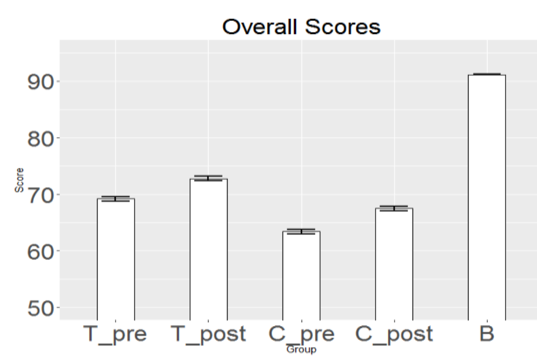

Figure3. Mean scores of tone productions in T3+T3 and T3+T1/T2/T4

3.1 Mandarin Third Tone Sandhi Rule

For the third tone sandhi rule (T3+T3), Cantonese speakers from TG obtained mean values of 78.45 in pre-test and 81.83 in post-test ($p < .001$), whereas CG insignificantly ($p = 0.3149$) decreased from the mean score of 69.72 to 68.52 in two tests . As shown in Figure 4, the scores for both pre- and post-test among TG and CG were significantly ($p < .001$) lower than Beijing speakers (M=89.99).

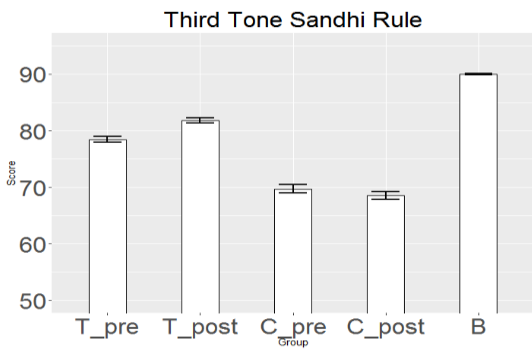

Figure4. Mean scores of tone productions in T3+T3

Trained Words and New words: Mean rating scores obtained from trained words and new words were respectively displayed in Figure 5 and Figure 6. Trained words referred to the three real words used to train TG in the training sessions, and new words did not appear in the training session. While for speakers of CG, they had congruent exposures to the “trained word” and “new words” due to their absence of perception training session.

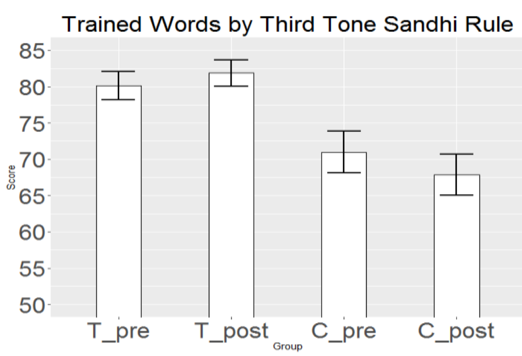

Figure5. Mean scores of trained words in T3+T3

Scores of TG on trained words (Figure 5) increased from 80.14 to 81.89 without significance ($p = 0.5151$), while the rating scores of new words (Figure 6) significantly increased from 78.33 to 81.82 ($p < .001$). The ratings of “trained words” only involved three real words (/məi3/ /xau3/ “美好”, /ʂʷəi3/ /kʷoo3/ “水果” and /ʈʂʷæn3/ /lʲæn3/ “转脸”) while “new words” had an overwhelming majority in their production. For CG, they decreased from 70.97 to 67.85 (p=0.44) and from 69.63 to 68.56 (p=0.30) respectively for “trained words” and “new words”.

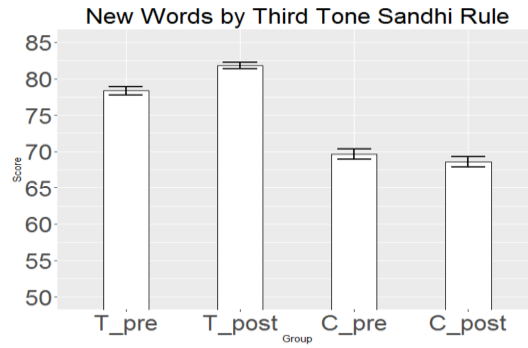

Figure6. Mean scores of new words in T3+T3

Real Words and Wug Words: Judgements of speech production by TG showed that the performances of both real and wug words were better in post-tests. In figure 7 and figure 8, the mean rating scores were 77.59 (pre) and 80.76 (post) for real words ($p < .001$), and 78.62 (pre) and 82.04 (post) for wug words ($p < .001$) for TG. However, evaluation results of CG decreased from 65.19 (pre) to 62.82 (post) in real words (p=0.3696), and 70.62 (pre) to 69.66 (post) in wug words (p=0.3728) of T3+T3 combination.

To investigate whether lexical meaning affected the application of tone sandhi rules, paired samples t-test was applied to compare the rating scores obtained from real words and wug words among TG, CG and Beijing speakers. Current results indicated that native Beijing speakers produced real words (M=90.15) and wug words (M=90.00) similarly in terms of third tone sandhi rule (T3+T3). This consistent performance in real and wug words also occurred in the production by Cantonese speakers of TG, while speakers of CG pronounced real words and wug words differently either in pre-test ($p = 0.008$) and post-test ($p = 0.008$).

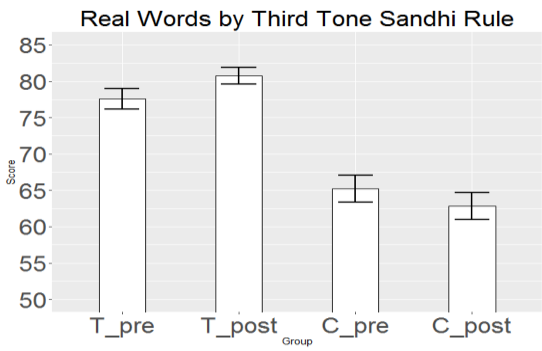

Figure7. Mean scores of real words in T3+T3

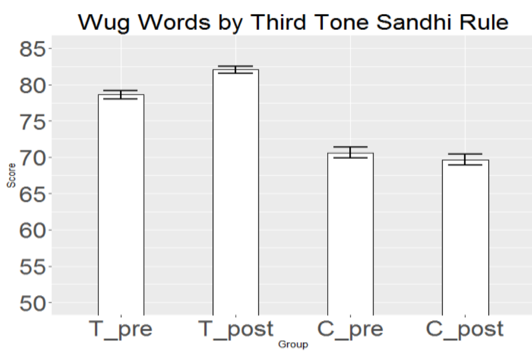

Figure8. Mean scores of wug words in T3+T3

3.2 Mandarin Half-Third Tone Sandhi Rule

Evaluation results of speech production with Mandarin half-third tone sandhi rule (T3+T1/T2/T4), as well as comparison with native Mandarin speakers, were generally demonstrated in Figure 9. Mean rating scores of TG speakers increased from 66.06 to 69.75 after training ($p < .001$). Mean scores obtained from CG showed greater improvement from pre-test (61.27) to post-test (67.15) ($p < .001$). Besides, rating scores of speech production by Cantonese speakers considerably inferior to that of Beijing speakers (M=91.56).

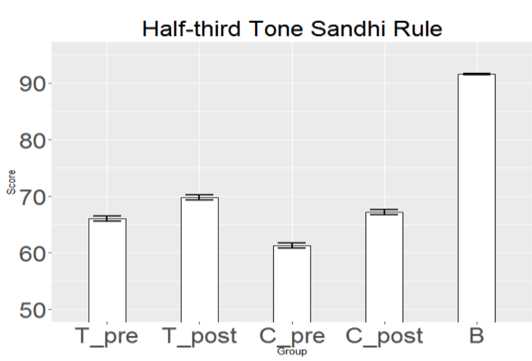

Figure9. Mean scores of tone productions in T3+T1/T2/T4

Trained Words and New Words: Figure 10 and Figure 11 showed mean rating scores of trained words and new words in term of the half-third sandhi rule (T3+T1/T2/T4) for TG and CG. Generally, performance of TG significantly improved in either trained words or new words. They made considerable increase of mean values from 67.79 in pre-test to 77.35 in post-test ($p = 0.0276$) for trained words (Figure 10.). At the same time, scores of new words production increased from 66.03 to 69.59 (Figure 11.) after training ($p < .001$). For CG, their mean scores were 70.93 (pre) and 72.54 (post) for "trained words" ($p=0.7186$) and 61.06 (pre) and 67.03 (post) for "new words" ($p<.001$).

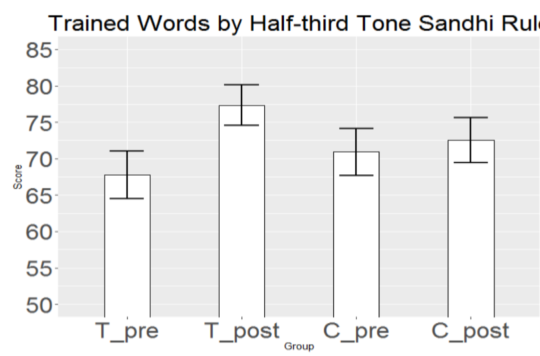

Figure10. Mean scores of trained words in T3+T1/T2/T4

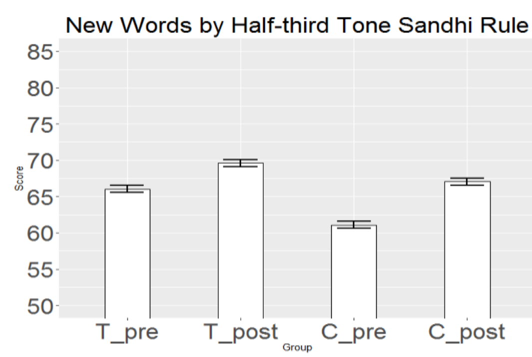

Figure11. Mean scores of new words in T3+T1/T2/T4

Real Words and Wug Words: Scores of real words and wug words applied with Mandarin half-third tone sandhi rule by TG and CG were visualized in Figure 12 and Figure 13. The mean scores of real words (Figure 12.) were 71.83 in the pre-test and 77.59 after training ($p < .001$). For wug words, they gained an average of 64.91 and 68.18 before and after training ($p < .001$). Similar to TG, performance of CG also improved and was manifested by scores from 69.81 to 77.50 for real words (p<.001) and from 59.56 to 65.08 for wug words (p<.001).

We also compared the effects of lexical meaning on speech production when applying the half-third sandhi rule (T3+T1/T2/T4). Consistent with native speakers, as Beijing speakers obtained 92.23 for real words and 91.42 for wug words with significant difference (p < .001), Cantonese speakers of TG and CG both significantly differed in producing real words and wug words. Specifically, mean scores of TG were 71.83 and 64.91 respectively for real and wug words (p < .001) in the pre-test. After training, the evaluation of real words (77.59) was still significantly different from that of wug words (68.18) (p < .001). Meanwhile, speakers of CG also produced better real words than wug words on perceptual evaluation perspective. They obtained mean rating scores of 69.81 for real words and 59.56 for wug words (p < .001) in the pre-test, as well as 77.50 for real words and 65.08 for wug words (p<.001) in the post-test.

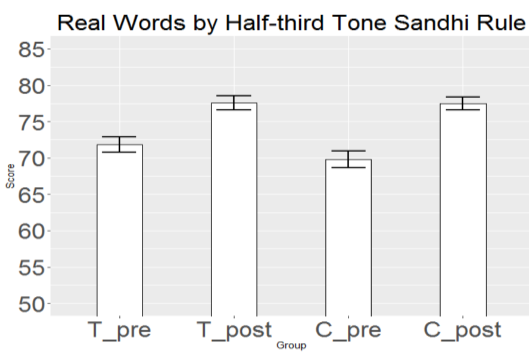

Figure12. Mean scores of real words in T3+T1/T2/T4

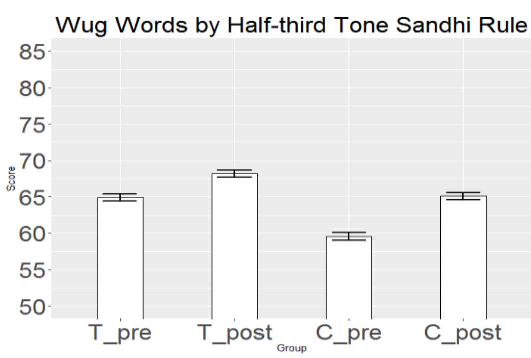

Figure13. Mean scores of wug words in T3+T1/T2/T4

4 Discussion and Conclusion

4.1 Third Tone Sandhi Rule

This pilot study of short-term laboratory training shows that the application of L2 Mandarin tone sandhi rules by Cantonese speakers is generally improved from pre-test to post-test after perceptual training. In T3+T3, trainees significantly improves their performance in the overwhelming majority of words (new words), including real words and wug words. Compared with the overall decrease of mean scores among speakers without training of CG, it is reasonable to maintain that perceptual training is effective for Cantonese speakers to apply the Mandarin third tone sandhi. Furthermore, the performance of speakers for control group rarely shows correlation to either familiarity with words or lexical meaning.

However, for Mandarin third sandhi rule (T3+T3), Cantonese speakers did not show significant improvement in trained words after perceptual training. One possible reason is that they have already produced the trained words well enough as they obtained quite high scores evaluated by native Mandarin raters in the pre-test. According to Wayland (2004), prior experience with native tone language may be transferable to the perception of tone in another non-native tonal language. Therefore, it is possible that some Cantonese speakers, as native speakers of a tonal language, have acquired the ability to modify or change the pitch couture of Tone 3 followed by another Tone 3 in Mandarin before attending the experiment. It is thus worthwhile to compare the results from the current study with the application of Mandarin tone sandhi rules by non-native speakers with a non-tonal language background (Li et al., 2018).

Besides, we also investigated their typical error patterns to explain why Cantonese speakers tend to behave similarly on the application of sandhi T3 in pre- and post-tests. One finding shows that the underlying sandhi T3 may be mispronounced as full T4 in Mandarin. For example, the AO+AG (Type3 in section 2.2) combination of /məi3/ /tʲəu3/ may be produced as /məi4/ /tʲəu2/ for some speakers. Because Cantonese speakers often produce the second T3 in T3+T3 context as Mandarin T2, a new context consisting of T3+T2 with half-third sandhi would be consequently produced. In this case, Cantonese speakers might easily apply the more phonetic motivated rule, T3 (213) → 21/_ {T1 (55), T2 (35), T4 (51)}, to change underlying sandhi T3 to T4 of Mandarin. Note that Cantonese speakers' misperception of T3 as T2 is caused by their perceptual assimilation of Mandarin tones. Another phenomenon is that Cantonese speakers probably produce two isolated Tone 3 without sandhi. In other words, they imitate the two monosyllables rather than produce them into a disyllabic word.

Additionally, there is no agreement on whether Cantonese speakers apply the third sandhi rule differently in pronouncing real words and wug words. On one hand, speakers of training groups pronounce real words and wug words without difference either before or after training. The acoustic study (Chen et al., 2017) on Mandarin production by Cantonese speakers suggests that Cantonese speakers are able to apply Mandarin third tone sandhi rule with a computational mechanism. In other words, Cantonese speakers tend to apply this rule on real words and wug words in the same way. Meanwhile, our present evaluation results show that there is no significant difference in real and wug words for Beijing speakers, which contradicts with Zhang and Lai (2010)'s findings that Beijing Mandarin speakers differently pronounce real and wug words in the context of Tone 3 followed by Tone 3. On the other hand, Cantonese speakers of the control group pronounced real words and wug words differently in both pre- and post-test. Therefore, it is still controversial whether Cantonese speakers can apply Mandarin third tone sandhi rule with a computational mechanism.

4.2 Half-third Tone Sandhi Rule

For Mandarin half-third sandhi rule, it is probable that native tonal language speakers can apply it due to higher familiarity with words, especially on real words. The possibility that Cantonese speakers apply the half-third sandhi rule better in post-test is realized by their prior knowledge of L1 tonal system. Specifically, Cantonese speakers tend to map Mandarin Tone 3 (214) into Cantonese T4 (21) according to the perceptual assimilation of Mandarin tones. Tone 3 (214) should be changed into "half of Tone 3" (21) in T3 + T1/T2/T4, which is easy for Cantonese speakers by producing a tone (21) with the same pitch height to their L1 tone T4 (21) of Cantonese. Considering the fact that speakers of control group demonstrate general progresses in post-test, not only can perceptual training facilitate the transfer to tone sandhi production, but repetitions also improve the accuracy and naturalness of speech production in T3 + T1/T2/T4 for Cantonese speakers.

Moreover, the visual input of f0 contour companied with auditory may play a role in the process of Mandarin half-third sandhi application. As the f0 contour of half-third sandhi makes a larger difference from that of the full Tone 3 in isolation, trainees are able to easily distinguish this sandhi from full Tone 3 and then process it. Whereas the f0 contour of the third tone sandhi is likely to be falling-raising.

Typical error patterns are also found in the application of half-third sandhi rule in present evaluation study. Challenges for Cantonese speakers might relate to the perceptual assimilation of T3 to T2 of Mandarin. Taking /məi3/ /lʲii4/ as an example, a number of speakers produce this disyllabic word as /məi2/ /lʲii4/ instead of /məi21/ /lʲii4/. It is likely that speakers tend to raise the offset of proceeding tone when followed by a tone with higher onset. However, underlying half-third tone in /məi3/ /kʰʷoo2/ is also mispronounced as Tone 2 like /məi2/ /kʰʷoo2/.

As to lexical meaning, Cantonese speakers might lexically apply the half-third tone sandhi rule. They differed in producing real words and wug words before and after training. Considerable improvements due to perception training prove that perceptual acquisition can be transferred to production with respect to this more phonetic motivated rule. However, it is necessary to conduct further long-term training and to investigate whether they can retain the improvement after a period of time.

In sum, the current perceptual evaluation is a pilot study to examine the production of Mandarin tone sandhi by non-native tonal speakers with and without perceptual training. From the perceptual perspective, Cantonese speakers are able to transfer their perceptual learning of Mandarin tone sandhi to immediate production evaluated by native listeners at certain degree. At the same time, high familiarity due to repetitions also increases the application degree of half-third sandhi rule to certain extent. Perceptual training is much more effective on the more phonetic motivated rule T3 (213) → 21/_ {T1 (55), T2 (35), T4 (51)}, especially on real words. As for third sandhi rule, present perceptual evaluation did not show dramatic improvement in speech production by Cantonese speakers. Our findings will be complementary to our further acoustic results of tone sandhi production by Cantonese speakers in pre- and post-tests.

Lastly, more raters, long-term perceptual training and investigation of individual variabilities are three important factors to be considered in future related studies.

Acknowledgements

This study is supported by grants from Faculty of Humanities [grant number: 1-ZVHH] and Department of Chinese and Bilingual Studies [grant number: 88C3], the Hong Kong Polytechnic University.

References

Antoniou, M. and Chin, J. L., 2018. What Can Lexical Tone Training Studies in Adults Tell Us about Tone Processing in Children? Frontiers in psychology, 9, p.1.

Bauer, R.S. and Benedict, P.K., 1997. Modern canton-ese phonology (Vol. 102). Walter de Gruyter.

Chao, Y. R., 1948. Mandarin primer: An intensive course in spoken Chinese. Harvard University Press.

Chen, A., Liu, L. and Kager, R., 2015. Cross-linguistic perception of Mandarin tone sandhi. Language Sciences, 48, pp.62-69.

Chen, M.Y., 2000. Tone sandhi: Patterns across Chinese dialects (Vol. 92). Cambridge University Press.

Chen, S., He, Y. J., Yuen, C.W., Li, B. and Yang, Y., 2017. Mechanisms of tone sandhi rule application by non-native speakers. Proc. Interspeech 2017, 1760-1764.

Cooper, A. and Wang, Y., 2013. Effects of tone training on Cantonese tone-word learning. The Journal of the Acoustical Society of America, 134(2), EL133-EL139.

Francis, A.L., Ciocca, V., Ma, L. and Fenn, K., 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. Journal of Phonetics, 36(2), pp.268-294.

Hao, Y.C., 2012. Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. Journal of phonetics, 40(2), pp.269-279.

Hao, Y.C. and de Jong, K., 2016. Imitation of second language sounds in relation to L2 perception and production. Journal of Phonetics, 54, pp.151-168.

Kaan, E., Wayland, R.P., Bao, M. and Barkley, C.M., 2007. Effects of native language and training on lexical tone perception: An event-related potential study. Brain research, 1148, pp.113-122.

Li, B., Yang Y. and Chen S., 2018. Evaluating the Effects of Perceptual Training on Application of Mandarin Tone Sandhi Rules by English Speakers. In Proceedings of the 2018 International Conference on Bilingual Learning and Teaching, Hong Kong.

Wang, Y., Spence, M.M., Jongman, A. and Sereno, J.A., 1999. Training American listeners to perceive Mandarin tones. The Journal of the acoustical society of America, 106(6), pp.3649-3658.

Wang, Y., Jongman, A. and Sereno, J.A., 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. The Journal of the Acoustical Society of America, 113(2), pp.1033-1043.

Wayland, R., 1997. Non-native production of Thai: Acoustic measurements and accentedness ratings. Applied linguistics, 18(3), pp.345-373.

Wayland, R. P. and Guion, S., 2003. Perceptual discrimination of Thai tones by naïve and experienced learners of Thai. Applied Psycholinguistics, 24(1), pp.113-129.

Wayland, R.P. and Guion, S.G., 2004. Training English and Chinese listeners to perceive Thai tones: A preliminary report. Language Learning, 54(4), pp.681-712.

Wayland, R.P. and Li, B., 2008. Effects of two training procedures in cross-language perception of tones. Journal of Phonetics, 36(2), pp. 250-267.

Zhang, J. and Lai, Y. W., 2010. Testing the role of phonetic knowledge in Mandarin tone sandhi. Phonology, 27(1), pp.153-201.

Zhang, J., Lai, Y.W. and Sailor, C., 2011. Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. Lingua, 121(2), pp.181-206.

Zhang, C., Xia, Q. and Peng, G., 2015. Mandarin third tone sandhi requires more effortful phonological encoding in speech production: Evidence from an ERP study. Journal of Neurolinguistics, 33, pp.149-162.