

# Carryover Effects on Tones in Hong Kong Cantonese

Bin LI<sup>1</sup>, Yihan GUAN<sup>1</sup>, Si CHEN<sup>2</sup>

<sup>1</sup>City University of Hong Kong

<sup>2</sup>The Hong Kong Polytechnic University

binli2@cityu.edu.hk, yihanguan2@cityu.edu.hk, sarah.chen@polyu.edu.hk

## Abstract

Tonal coarticulation has been found in connected speech in many tonal languages. The current study investigated the carryover effect on tones of Hong Kong Cantonese (Cantonese), by examining F0 perturbation and preservation of two rising tones (T2 and T5) under the influence of a preceding tone (Tx). Twenty-three young adults were recruited to read a wordlist containing disyllabic words (Tx+T2/T5) and monosyllabic characters in six lexical tones. Onsets, offsets, and contours of T2 and T5 were measured and entered for statistical analyses. The carryover effect of high and low tones was found on T2 and T5. The coarticulation triggered by the low tone carried throughout the entire tonal contour while the coarticulation triggered by the high tone did not, as offsets of pitch contours of T2 and T5 preceded by a high tone reached comparable levels of their own underlying pitch targets.

**Index Terms:** tonal coarticulation, carryover effects, Cantonese tones, F0 perturbation and preservation

## 1. Introduction

Lexical tones are known to be subject to coarticulatory effects imposed by adjacent tones in connected speech. Such coarticulatory influence may take place in both directions, that is, carryover and anticipatory. The universal properties of tonal coarticulation are found across tonal languages including Thai [1], Vietnamese [2], Mandarin [3], Taiwanese [4], Yoruba [5] and some other West African languages [6]. Both carryover and anticipatory effects have been found in these studies, but the carryover effect appears to be stronger than the anticipatory effect in many languages [1, 2, 3], i.e. the number of tonal categories and contexts the carryover effect can influence, as well as the temporal extent of the effect, are greater than those correlated with the anticipatory effect, which is constrained by certain suprasegmental contexts, e.g. preceded by high-pitch offsets [1, 3]. Exceptions like Nanjing Chinese [7] and Malaysian Hokkein [8] were also reported, where comparable applicability and strength of both carryover and anticipatory effects have been found.

In terms of their directionality, the carryover effect tends to be assimilatory [1, 3], that is, a tone with a high-pitch offset raises the onset pitch of the following tone, while the anticipatory effect may show more patterns [3, 7] where a preceding tone's offset rises due to anticipating low-pitch offset of its following tone.

The current study examines the carryover effect in Hong Kong Cantonese (Cantonese), where there are six lexical tones [9], including four level tones and two rising tones, as shown in Table 1 and Figure 1 (based on data extracted from recordings of monosyllables read by speakers in this study).

Table 1: Cantonese tones in Chao's notation

| Tone | Name             | 5-point notation |
|------|------------------|------------------|
| 1    | High level tone  | 55               |
| 2    | High rising tone | 25               |
| 3    | Mid level tone   | 33               |
| 4    | Low falling tone | 21               |
| 5    | Low rising tone  | 23               |
| 6    | Low level tone   | 22               |

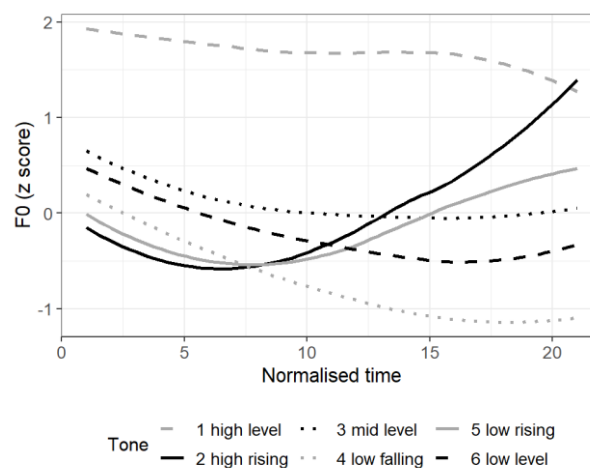


Figure 1: Normalised F0 contours of six Cantonese lexical tones.

The normalised F0 contours illustrated in Figure 1 show that T1 and T2 have relatively high offset F0, while T4 ends extremely low and can be creaky [10]. T2 and T5 are both rising, but they reach distinct levels at offset, which serves as the main cue for tone distinction in production and perception. Research on tone coarticulation in Cantonese [11], however, is scarce, which reports that carryover effect manifest in the first half of the tone bearing syllable, which serves for the tonal transition, while the tonal contours in the second half of the syllable remain relatively distinct. This is also interpreted as support to the reliance of the latter portion of syllables in tone identification in connected speech [12]. However, given the small number of speakers (four in [11]) and lack of statistical validation, the carryover effect identified in previous research calls for re-examination. This study intends to answer the following questions:

- Is there any carryover effect in Cantonese tones?
- How far does the effect sustain in a tonal syllable? Will it last till the tonal offset?

c) Do high and low-pitch offsets have the same or different effects on the following tones?

## 2. Method

### 2.1. Subjects

Twenty-three native speakers of Hong Kong Cantonese (19 born and raised in Hong Kong, and 4 born in Cantonese-speaking areas in the mainland and moved to Hong Kong before 12) were recruited on a voluntary basis and recorded by students in an undergraduate linguistics course. Ten speakers were male, and the rest 12 were female. Their age ranged from 19-25 ( $Mean = 21$ ,  $SD = 2.06$ ). They all gave written consent to participate in the recording and provide ethnographic information for research purposes.

### 2.2. Stimuli

Twelve disyllabic words were designed to test the carryover effect. Tones of the two syllables in each word were Tx+T2/T5, where Tx was one of the six lexical tones in Cantonese and T2 or T5 was used in the second syllable. In each minimal pair, segmental elements in both syllables were kept identical. For example, in T1+T2 and T1+T5 pair, the first syllable was *syu* bearing T1 and the second syllable was *se* with either T2 or T5. All words were real and meaningful. The two rising tones were selected to examine whether different offset values of a rising tone would be influenced differently by the carryover effect from the preceding tones. Words used in the study are listed in Table 2. In addition to the disyllabic words, six monosyllabic characters with all lexical tones in Cantonese were used to elicit reading in citation tones: si1 ‘poem’ 詩, si2 ‘history’ 史, si3 ‘try’ 試, si4 ‘time’ 時, si5 ‘city’ 市, si6 ‘is/yes’ 是.

Table 2: List of disyllabic words for tone production.

| Tx + T2 |         |              |                               |
|---------|---------|--------------|-------------------------------|
| Tx      | Chinese | Jyutping     | Gloss                         |
| 1       | 書寫      | syu1 se2     | write                         |
| 2       | 古史      | gu2 si2      | ancient history               |
| 3       | 國史      | gwok3 si2    | national history              |
| 4       | 虹口      | hung4 hau2   | Hongkou (a place in Shanghai) |
| 5       | 老虎      | lou5 fu2     | tiger                         |
| 6       | 幻想      | waan6 soeng2 | Fantasy                       |
| Tx + T5 |         |              |                               |
| Tx      | Chinese | Jyutping     | Gloss                         |
| 1       | 書社      | syu1 se5     | publishing house              |
| 2       | 股市      | gu2 si5      | stock market                  |
| 3       | 國試      | gwok3 si5    | national exam                 |
| 4       | 雄厚      | hung4 hau5   | robust                        |
| 5       | 老婦      | lou5 fu5     | old woman                     |
| 6       | 患上      | waan6 soeng5 | suffer from (an illness)      |

Stimuli were presented in traditional Chinese characters and randomised to avoid priming effects [13]. Each speaker was asked to read the list twice in a quiet place. Recordings were made using mobile phones. Manual check of mobile phone’s specification and sample quality was conducted to ensure that recordings were good and valid for the subsequent analysis.

### 2.3. Data processing

Manual checking was carried out to validate the recorded data, making sure that tokens were correctly produced and their f0 contours were clear. Excluding those with wrong pronunciation or with bad quality, a total of 1,260 syllables were obtained and used in the subsequent analysis. Segmentation and labelling were completed manually. The F0 values were extracted from 21 equidistant points in the voicing part of each syllable, using Prosody Pro [14] in Praat [15]. The extracted F0 values were within 75 Hz and 600 Hz, which were all checked manually to ensure the first and the last points were the beginning and end of periodicity within each extracted F0 contour. When creaky voice was noticed, which was frequent in the extremely low T4, the contours were corrected using the Praat manipulation editor, or excluded if correction was not able to achieve. Each corrected pitch value was manually extracted.

To facilitate acoustic analysis of all speakers, after F0 points extraction, z normalisation was done using speaker-specific mean and standard deviation. All 21 points were entered in normalisation formula, following studies on tonal coarticulation, such as [7].

## 3. Results

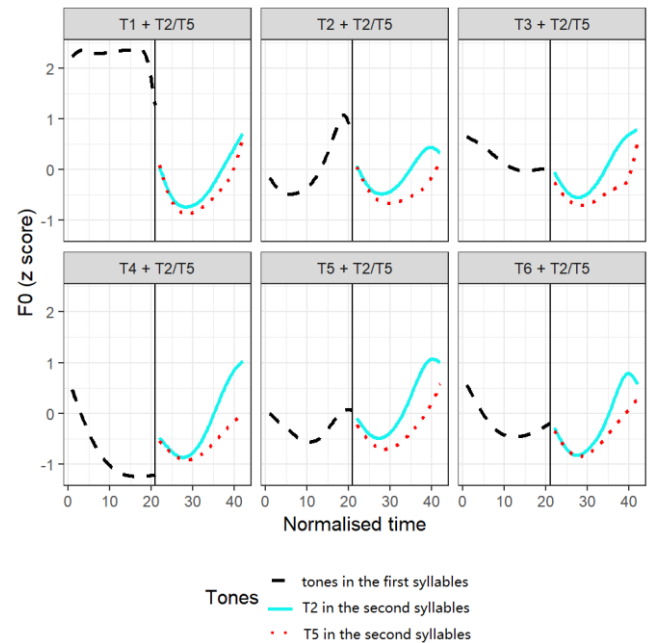


Figure 2: Normalised F0 contours of Tx + T2/T5 drawn using gam smooth method. In each panel: the vertical line at centre indicates the syllable boundary within a word; dashed curves represent Tx in the first syllable, solid curves are T2 and dotted curves T5.

Figure 2 presents the tonal contours from all twelve disyllabic words containing Tx+T2/T5. Statistical analysis was carried out to examine if the carryover effect was present and if there was any pattern manifested. Linear mixed-effects regression models were built using the *lme4* package [16] in R [17], where p-values were obtained by the *lmerTest* package [18]. Tonal contours were compared in a generalised additive mixed model built using the *mgcv* package [19], when significance at onset and offset was shown, and contours were

plotted using the *itsadug* package [20]. The results are presented and discussed as follows.

### 3.1. Carryover effects on F0 onset

With the F0 values extracted from disyllabic words (Tx+T2/T5), we built a linear mixed-effects regression model to examine whether preceding tones have any carryover effect on the following tones. Z scores of F0 at the onset of T2/T5 were treated as the dependent variable. Two rising tones, the neighbouring tones, as well as their interactions were the independent variables. In the models, the onset z scores of T2 were treated as the reference. Repetition of stimuli was entered as a control variable, and speakers were entered as the random variable.

First of all, no significant difference is found between the onset of T2 and T5 ( $b = 0.07, t = 0.56, p = 0.57$ ). As T2 is set as the reference, the result is interpreted as that onsets of the two rising tones are similar. T2 and T5 preceded by T1 both have a significantly higher onset ( $b = 0.26, t = 2.10, p < 0.05$ ) in the surface contours than in their underlying pitch target, and those preceded by T4 (low falling) have a significantly lower onset ( $b = -0.38, t = -3.11, p < 0.01$ ). When T2 and T5 are preceded by T2, T3, T5 and T6, there is no significant difference from their pitch targets as shown in the monosyllabic reading in terms of F0 onsets (all  $|b/s| < 0.26, |t/s| < 1.81, ps > 0.07$ ). In terms of their interaction, when T5 is preceded by T3 (mid-level), T5 shows a significantly lower F0 onset ( $b = -0.35, t = -2.00, p < 0.05$ ). Neither repetition in production nor any other interaction terms between T5 and neighbouring tones contributes to F0 perturbation at onset (all  $|b/s| < 0.18, |t/s| < 1.07, ps > 0.54$ ).

In sum, the results above show that a high-pitch offset, e.g. that of T1, and a low-pitch offset, e.g. that of T4, affect the onset of following tones. The carryover effects are assimilatory, which is present in the increase and decrease of F0 onsets respectively in the following tones.

### 3.2. Carryover effects on F0 offset

The previous study [13] finds that the carryover effect usually occurs in the first half of a tone-bearing syllable. As we found clear and strong carryover effects present as early as at onset, next, we examine the temporal extent the carryover effect would last and whether they would have different influences on different offsets with different pitch height. Two linear mixed-effects regression models were built: one for T2 and the other for T5. Dependent variables in two models were Z scores of F0 offsets, and the independent variables were preceding tones, where the citation form was treated as with a virtual null preceding tone serving as the reference. Repetition of production was entered as a control variable and speakers were entered as the random variable in the model.

First, F0 offsets of T2 when preceded by six tones, are all significantly lower than those of its pitch targets in citation forms (all  $|b/s| > 0.87, |t/s| > 5.77, ps < 0.001$ ). Repetition does not affect offset variation ( $b = 0.07, t = 0.94, p = 0.35$ ).

F0 offsets of T5 are significantly lower when preceded by T4 (low falling) than those of its pitch targets in citation forms ( $b = -0.50, t = -2.86, p < 0.01$ ). None of the rest preceding tones, i.e. T1, T2, T3, T5 and T6, triggers perturbation in F0 offsets of T5 (all  $|b/s| < 0.28, |t/s| < 1.75, ps > 0.08$ ). Repetition does not contribute to variation in F0 offsets either ( $b = -0.11, t = -1.30, p = 0.19$ ).

T2 and T5 both rise from low, but end at a high and mid offset respectively. The significant difference at F0 offsets of T2 and T5, triggered by a preceding T4, is expected and in conformity to the lowering onsets by the same tone T4 that claims the lowest offset in the Cantonese tonal space. F0 perturbation at both onset and offset suggests that the carryover effect sustains throughout the tonal part of T5. When T5 is preceded by the high-level T1, it shows a relatively higher onset than its pitch target in the citation form, but the rise soon reduces to a level similar to the citation. The perturbation patterns at onsets and offsets of T2 and T5 suggests distinct coarticulation influences by low tones and high tones in Cantonese. The carryover effect triggered by a low tone, T4, is sustained from the beginning to the end of the following tone, whereas the effect by a high tone, T1, is significant at onsets but diminish at offsets of the following tone.

In sum, we found that T4, the low falling tone has exerted the biggest effects on the following low rising T5, where the effects appeared at the offsets of T5 in the second syllable.

### 3.3. Carryover effects on the tone contour

In order to see whether carryover effects exerted by T4 on the subsequent T5 would cover its whole contour, a generalised additive mixed model (GAMM) was built, as GAMMs are suitable to analyse data with non-linear relationships. In the model, the contours of the citation T5 were treated as the reference smooth and then T5 preceded by T4 as difference smooths. Speakers were entered as the random intercept, and between-speaker variation developed over time were treated as the random smooth in the model.

The model summary shows significant differences between the two conditions in terms of F0 range (for parametric terms:  $b = -0.45, std.Error = 0.05, t = -22.11, p < 0.001$ ). The smooth term, however, indicates that the two smooths are similar in shape (for smooth terms:  $edf = 1.55, Ref.df = 1.92, F = 0.65, p = 0.44$ ), as illustrated in Figure 3. As pointed out in [21], the GAMM outputs alone cannot sufficiently provide comprehensive interpretation of the results, and therefore, visual inspection of F0 tracks is necessary. We can thus conclude from Figure 3 below that the carryover effect triggered by T4 sustains till the end of T5 without changing the shape of its contour.

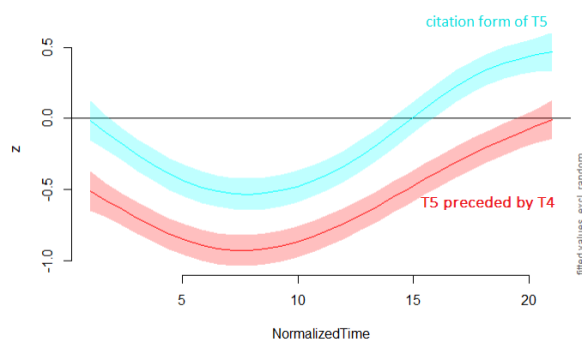


Figure 3: Visualisation of T5 F0 contours in citation form (upper curve) and when preceded by T4 (lower curve). The 95% confidence intervals are shown by shaded bands.

## 4. Discussion

Data analysis above reveals an interesting phenomenon in Cantonese tonal coarticulation, that offsets of T2 are always lowered by its preceding tones, no matter high, mid or low. T2 and T5 have been reported to gradually merge in recent decades among the young generation [22-25], despite the different perspectives to how the tone merger develops, e.g. the offset of T5 approaches the offset target as T2 [22], or the offset of T2 is lowered to meet that of T5 [23], or either is possible among different groups of speakers [24], or both tones are dynamically approaching each other [25]. F0 perturbation of T2 observed here could thus reflect the on-going sound change in Hong Kong Cantonese, which confirms that the overall lowering effects on T2 brought by the preceding tones can be an indication that mergers may happen more often from T2 where the offsets are lowered in connected speech where full articulation of tones is not common. However, there is no clear evidence to show that carryover effects are related to the rising tone merger in Cantonese. Future studies can follow this direction to investigate whether the offset lowering in T2 is due to downdrift [26] since 5 in Chao's notation is relatively hard to reach in connected speech and can be realised hypoarticulately, whether coarticulation is related to the tonal change, and also whether the current tone height is sufficient for the perceptual distinction between T2 and T5.

In the combination T3+T5 (33+23 in Chao's notation) where two tones are within the mid-range of tonal space, the onset of T5 is found lowered but not the offset. There are two possibilities in the cause of this process. First, speakers complete F0 transition before the offset of T5, where the offset is then exempted from the carryover effect. Secondly, in the disyllabic combination where tonal space is so packed, speakers may need over articulate the rising contour of T5 by lowering its onset in order to distinguish it from T3 in the same F0 range. Therefore, if speech planning is considered, it is difficult to separate the unintended and intrinsic phonetic effects from intended and active language mechanisms. Similarly, it may appear premature to conclude that the onset lowering of T5 when preceded by T3, is due to the carryover effect.

## 5. Conclusions

The current study aims to examine tonal coarticulation in Cantonese, with a focus on two rising tones. We set out to answer questions regarding the carryover effect in disyllabic words and our findings are as follows.

- a) There is tonal coarticulation in terms of the carryover effect among Cantonese tones in connected speech.
- b) The carryover effect may influence the offset of the following tone.
- c) High and low-pitch offsets have different effects on the following tones.

In sum, this study confirms that tonal effects in tonal coarticulation. i.e. a low tone exerts stronger carryover effects on T5, the low rising tone, while a high tone does not carry such strong effect or influence with a sustained temporal extent. F0 offsets of tones preceded by a high tone resume soon after onset lowering to the pitch level of its citation reading. Our findings provide empirical data and evidence to Cantonese tones in connected speech, which contributes to an updated typology of tonal coarticulation properties. Further investigation is called for to identify effects of tone merger and thus to confirm on the true mechanism underlying tonal coarticulation in Cantonese.

## 6. Acknowledgements

This study is supported by EDB(LE)/P&R/EL/175/12 and CityU TDG 6000653. We thank students in LT2203 (Language and Society) for their contribution in data collection.

## 7. References

- [1] J. Gandour, "Tonal coarticulation in Thai," *Journal of Phonetics*, vol. 22, pp. 477-492, 1994.
- [2] M. S. Han, K. O. Kim, "Phonetic variation of Vietnamese tones in disyllabic utterances," *Journal of Phonetics* vol. 2, pp. 223-232, 1974.
- [3] Y. Xu, "Contextual tonal variations in Mandarin," *Journal of Phonetics*, vol. 25, pp. 61-83, 1997.
- [4] S.-h. Peng, "Production and perception of Taiwanese tones in different tonal and prosodic contexts," *Journal of Phonetics*, vol. 25, no.3, pp. 371-400, 1997.
- [5] Y. O. Laniran, "Intonation in tone languages: The phonetic implementation of tones in Yoruba". PhD dissertation, Univ Cornell. Lea, W.A. 1992.
- [6] L. M. Hyman and G. S. Russell, "Universals of tone rules: Evidence from West Africa," *Linguistic Inquiry*, vol. 5, pp. 81-115, 1974.
- [7] S. Chen, C. Wiltshire and B. Li, "An updated typology of tonal coarticulation properties," *Taiwan Journal of Linguistics*, vol. 16, no. 2, pp. 79-114, 2018.
- [8] Y. -h. Chang and F. -f. Hsieh, "Tonal coarticulation in Malaysian Hokkien: A typological anomaly?," *The Linguistic Review*, vol. 29, pp. 37-73, 2012.
- [9] S. Matthews and V. Yip, *Cantonese: A Comprehensive Grammar*, London: Routledge, 1994.
- [10] K. M. Yu, "The role of creaky voice in Cantonese tonal perception," *Journal of the Acoustical Society of America*, vol. 136, pp. 1320-1333, 2014.
- [11] Y. W. Wong, "Contextual tonal variations and pitch targets in Cantonese," *Proceedings of Speech Prosody 2006*, pp. 317-320, 2006.
- [12] Y. W. Wong, "Temporal distribution of tonal information in Continuous Cantonese speech," *Proceedings of International Symposium on Tonal Aspects of Languages 2006*, pp. 118-121, 2006.
- [13] L. Clark, "Priming as a motivating factor in sociophonetic variation and change," *Topics in Cognitive Science*, vol.10, pp. 729-744, 2018.
- [14] Y. Xu, "ProsodyPro — A tool for large-scale systematic prosody analysis," *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, pp. 7-10, 2013.
- [15] P. Boersma and D. Weenink, Praat: doing phonetics by computer. 2018, Ver 6.0.43. [Online] Available: <http://www.praat.org/>.
- [16] D. Bates, M. Maechler and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no.1, pp. 1-48, 2015.
- [17] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2013. [Online]. Available: <https://www.R-project.org/>.
- [18] A. Kuznetsova, P. B. Brockhoff and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *Journal of Statistical Software*, vol. 82, no.13, pp. 1-26, 2017.
- [19] S. Wood, *Generalized Additive Models: An Introduction with R*, 2<sup>nd</sup> edition. Chapman and Hall/CRC, 2017.
- [20] J. van Rij, M. Wieing, R. Baayen and H. van Rijn, "itsadug: Interpreting time series and autocorrelated data using GAMMs," R package ver 2.3, 2017.
- [21] M. Wieling, "Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English", *Journal of Phonetics*, vol. 70, pp. 86-116, 2018.
- [22] 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," *Language Variation and Change*, vol. 25, pp. 341-370, 2013.

- [23] B. Li and Y. Guan, "Generational differences in production of a tonal contrast in Hong Kong Cantonese," *Proceedings of the 19th International Congress of Phonetic Sciences*, pp. 186-190, 2019.
- [24] R. S. Bauer, K. H. Cheung and P. M. Cheung, "Variation and merger of the rising tones in Hong Kong Cantonese," *Language Variation and Change*, vol. 15, pp. 211-225, 2003.
- [25] R. S. Y. Fung and C. S. P. Wong, "Acoustic analysis of the new rising tone in Hong Kong Cantonese," *Proceedings of the 17th International Congress of Phonetic Science*, pp. 715-718, 2011.
- [26] P. C. M. Wong, "The effect of downdrift in the production and perception of CANTONESE level tones," *Proceedings of the 14th International Congress of Phonetic Sciences*, pp. 2395-2398, 1999.