9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

An Acoustic Study on the Effect of the Interaction between Intonation and Lexical Tones on the Realization of Cantonese Sentence-final Particles

Rachel Ka-Ying Tsui, Shelley Xiuli Tong

Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong
rachelkt@hku.hk, xltong@hku.hk

Abstract

Hong Kong Cantonese provides a fascinating case to the study of prosody because of its rich inventory of sentence-final particles (SFPs) that carry two levels of prosodic information: an inherent lexical tone and intonation. The present study examined the effect of the interaction between tone and intonation on the production of SFPs among native Cantonese speakers. Acoustic analyses were carried out to measure and compare: (1) interrogative and declarative SFPs that have the same underlying lexical tone, and (2) SFPs with their homophones. Results showed that there is an interaction between intonation and lexical tones within SFPs, such that the intonation contrasting declarative and interrogative sentences has an impact on the realization of the underlying lexical tones of the SFPs. However, such interaction is mainly caused by global pitch raising, rather than local pitch raising, signaling the interrogative status of the sentence.

Index Terms: Sentence-final particles, Cantonese, lexical tones, intonation, speech production

1. Introduction

Hong Kong Cantonese provides a unique window into the study of prosody, not only because Cantonese employs both tone and intonation, but also because it has a particularly rich inventory of sentence-final particles (SFPs) that carry both prosodic information: an inherent lexical tone and intonation. However, the question of how SFPs function in the prosodic system of a tonal language remains inadequately explored. The present study focuses on exploring the interaction between intonation and lexical tones through the rich system of Cantonese SFPs. While previous studies focused on how the presence of SFPs affects the intonation of the sentence to which they are attached, the present study extends previous research by examining the effect of the interaction between tone and intonation on the realization of SFPs. Through a speech production experiment, the present study explored to what extent SFPs fulfil the function at both the lexical prosodic level and the intonation level in Cantonese. Such an investigation would further inform us how the functions of intonation and lexical tones are distributed within a tonal language like Cantonese.

1.1. Lexical tones and intonation

Cantonese has a complex tonal inventory. There are six contrastive lexical tones, differing in terms of pitch height (high, mid, and low) and pitch contour (level, falling, and rising) [1, 2, 3]. A monosyllable can represent six different meanings in Cantonese with a change in tone only; for example, /si1/ (*poem*), /si2/ (*history*), /si3/ (*attempt*), /si4/ (*time*), /si5/ (*city*), and /si6/ (*verb-to-be*). Tone, therefore, plays a prominent role in distinguishing word meanings.

Meanwhile, such a rich inventory of tones in Cantonese poses a challenge to the study of prosody at a higher structural level, namely sentential intonation. Specifically, lexical tones and intonation are both manifested in changes in the fundamental frequency ($f0$) simultaneously [4, 5]. Many studies have suggested that, due to its complex lexical tone system, the extent to which intonation can be applied in Cantonese is restricted compared to non-tone languages like English [6, 7, 8, 9, 10, 11, 12, 13, 14]. In particular, [9] found that the difference in pitch contour between statements and questions was not found in the initial and medial positions of sentences; significant differences emerge only at the final syllable. In other words, the major location for the realization of intonation in Cantonese is restricted to the last syllable of a sentence, and do not expand to the preceding syllables within the same utterance [6, 9, 14, 15, 16, 17, 18].

1.2. Sentence-final particles

To compensate for the limited phonological space left for the realization of intonation, a rich inventory of SFPs has been developed to fulfil the intonation functions in Cantonese [17, 19, 20, 21]. SFPs refer to a class of bound grammatical words that attaches to the end of sentences to indicate sentential connotations, such as declarative statements and interrogative questions [11, 17, 20]. Each SFP is a monosyllable which is made of the segmental component (i.e., consonant and vowel) and an inherent lexical tone [17]. A considerable body of cross-language studies suggests that Cantonese SFPs express meanings that are usually conveyed through intonation in other languages (e.g. [11, 12, 17, 21, 22, 23, 24]).

The question of how SFPs are distributed and function in a within-language context, however, remains inadequately explored. Most studies characterizing the functions of SFPs are impressionistic, which were based on the auditory judgement of the authors who, most of the time, happen to be a native speaker of Cantonese [6, 20, 17]. Such kind of judgmental studies may not be able to fully capture the function of SFPs and their interaction with intonation in the prosodic system.

A limited number of studies to date have probed the role of SFPs in the prosodic system of Cantonese with systemic experimental and acoustic data (e.g., [14, 25, 26]). These studies targeted at investigating whether the presence of a SFP would restrict the occurrence of changes happening on the final lexical word within the sentence. These studies all found that final pitch rising for questions occur only in the final lexical words of those sentences without SFPs, but not in those preceding a SFP. In other words, the occurrence of pitch changes signaling the interrogative intonation appeared only in sentences without a SFP. The presence of the SFPs therefore seems to have eased the burden imposed by intonation on the final lexical word of the sentence.

10.21437/SpeechProsody.2018-9

However, none of these studies has addressed the issue regarding the interaction between intonation and lexical tones within SFPs. It therefore remains unclear whether intonation would impose additional intonational impact on the inherent tone of SFPs.

1.3. The present study

The aim of the present study is to further explore the interaction between intonation and lexical tones at the sentence-final position through SFPs. In particular, we explored whether lexical tones alone would be enough to fulfil the intonation functions—such that intonation will not exert any effect on the underlying tones of SFPs. Two sets of SFPs (one set indicating declarative function and another indicating interrogative function) were selected and acoustic analyses were conducted to explore the effect of the contrast between declarative and interrogative intonation on the production of SFPs bearing the same underlying lexical tone.

2. Method

2.1. Participants

Twenty-one native speakers of Cantonese were recruited (9 males, mean age = 20.62 years, *SD* = 22 months) from the University of Hong Kong. According to a language background questionnaire, Cantonese was reported to be the dominant and first language of all participants. None of them had a reported history of speech, hearing or language problems.

2.2. Stimuli and design

Two declarative SFPs and two interrogative SFPs were selected (see Table 1). The four target SFPs were used to form three experimental conditions: (1) SFPs being attached to the end of contextual sentences that denote a situation or event; (2) SFPs being placed in the middle of a neutral sentence; (3) homophones of the SFPs in the same sentence as in (2). Each speaker produced a total of 36 different utterances (= 4 targets words × 3 conditions × 3 repetitions).

Table 1: *The four target SFPs.*

SFPs	Connotations	Meaning
囉 /lo1/	Declarative	Pointing out the obviousness
啊 /a3/	Declarative	Softener of a statement
咩 /mɛ1/	Interrogative	Querying the truth
嘛 /ma3/	Interrogative	A neutral question

2.2.1. Condition 1 (SFP at the end of a contextual sentence)

Three sentences were designed: (1) /ka1 tsɛ1 tsɪn1 pan1 kʰɪk1/ (“*elder sister makes pancakes*”); (2) /ku1 tsɛ1 tsɔŋ1 tan1 sɪk1/ (“*aunt puts up string lights*”); (3) /sam1 sʊk1 pou1 kɐi1 tsʊk1/ (“*third uncle makes chicken porridge*”). Each sentence consisted of five syllables (a disyllabic subject + a monosyllabic verb + a disyllabic object). The use of the same syntactic structure can ensure they all bear the same prosodic structure.

With these three sentences, three experimental versions were formed. Each version was formed with two of the three sentences: Version 1 with sentence (1) and (2), Version 2 with (2) and (3), and Version 3 with (3) and (1). Within each version, each sentence was associated with the SFPs of one tone only, either T1 or T3. Participants were then asked to produce each of the two sentences twice: first time with the declarative SFP attached to the end of the sentence, and second time with the interrogative SFP.

2.2.2. Condition 2 (SFP in the middle of sentence)

A neutral sentence was used: /ŋɔ5 hɐi6 sœŋ2 nei2 tʊk6 ____ tsi6 sɪn5/ (“*I want you to first read the word ____*”). Each of the four target SFPs was inserted in the middle of the sentence to elicit the underlying tonal target of the SFPs (i.e., without the effect of sentence-final intonation).

2.2.3. Condition 3 (homophone in the middle of sentence)

The same carrier phrase in Condition 2 was used, with the (near) homophones of the four SFPs inserted to elicit their underlying tonal targets. As it was impossible to have homophones for all SFPs, near homophones were used in some cases (see Table 2).

Table 2: *The four (near-)homophones.*

Declarative SFP	Homophone	Interrogative SFP	Homophone
/lo1/	/wo1/; “*pot*”	/mɛ1/	/pɛ1/; “*teddy bear*”
/a3/	/a3/; “*second*”	/ma3/	/a3/; “*second*”

2.3. Procedure

The participants were randomly assigned to one of the aforementioned three experimental versions, with 7 participants for each. Recording was conducted individually in a sound-attenuated room located in the Division of Speech and Hearing Sciences of the University of Hong Kong. Recordings were recorded at a sampling rate of 44.1 kHz, with an AKG SE300 B microphone pre-amplifier, a CK91 cardioid condenser microphone capsule, and an Edirol USB Audio Capture UA-25, all of which were connected to a MacBook Pro computer.

All recordings were recorded via Praat [27]. A Praat script was run to present the target sentence on the Demo window. The participants were instructed to sit in front of a Philips 220S2 monitor, where the Demo window appeared and which was connected to the MacBook Pro computer running the recording program. The presentation of the target sentences was randomized across conditions and participants.

2.4. Data analysis

Praat [27] was used to perform acoustic analysis on the recordings obtained. Since the effect of intonation and lexical tone is mainly manifested in pitch patterns [28, 29, 30], measurements of pitch (i.e., fundamental frequency, *f*0) were collected for each target word across all three conditions and each word in the carrier sentences in Condition 1.

Spectrographic analyses were first carried manually to visually identify the voiced segment of each word from wide-band spectrogram and amplitude waveform displays. The onset of the voiced segment was manually marked at the first upward-going zero-crossing of the full vocal cycle of (1) the vowel for plosive-initial syllables, and (2) the whole syllable for sonorant-initial syllables. The offset of the voiced segment was marked by the loss of formants in the spectrographic display and periodic vocal cycle of the vowel [31, 32].

Fundamental frequency (*f*0) were then estimated, using the pitch profile generated by the autocorrelation algorithm in Praat [27], with the range of the *f*0 set between 75-600 Hz for males and 100-600 Hz for females. *F*0 was traced at nine evenly-spaced time points from the onset to the offset of the voiced segment of each target word [9, 31].

The *f*0 estimates of the nine time points were first obtained in Hertz (Hz) for producing graphical analyses, and were then converted into ERB rate (*equivalent rectangular bandwidth* rate) [33]. The ERB scale has been shown to correlate strongly with the frequency selectivity of our auditory system [34].

Averages of the pitch estimates of each target word at each of the nine time points were calculated for each of the condition for each speaker. The pitch contours of the target syllable, defined as the difference between the pitch offset and the pitch onset being divided by the duration of the syllable, were also calculated to indicate the pitch changes within the syllable. Group averages of these pitch measurements were then computed and submitted for later statistical analyses.

3. Results

3.1. Comparison of the target words across conditions

To explore if differences existed across the three experimental versions in Condition 1, a 3 (*version*: 1, 2, 3) × 2 (*tone*: T1 and T3) × 2 (*intonation*: declarative vs. interrogative) × 5 (*syllable*) ANOVA was conducted on the average of the nine time points for the five syllables in the carrier phrases. No significant effects nor interactions were found (*ps* > .100). The data obtained were collapsed in the following analyses.

The pitch slope values in ERB-rate for the four SFPs in Condition 1, 2, and 3 were submitted to a linear mixed effects model using R [35] and the *lme4* package [36]. The intercept for *subjects* and *genders* were entered in the random-effects structure. *Intonation* (declarative vs. interrogative), *tone* (T1, T3), and *condition* (Condition 1, 2, 3) were entered as fixed effects. Significance of the effects in question were obtained by the ANOVA function in the *car* package [37].

The effect of *intonation*, $\chi^2(1) = 14.03$, $p < .001$, and *tone*, $\chi^2(1) = 22.81$, $p < .001$, were found significant. The effect of *condition* was, however, not significant, $\chi^2(2) = .76$, $p = .685$. Yet, the linear mixed effects model revealed a significant interaction between *condition* and *intonation*, $\chi^2(2) = 46.88$, $p < .001$, as well as among *condition*, *intonation*, and *tone*, $\chi^2(2) = 19.61$, $p < .001$. Post-hoc tests with Tukey-adjusted comparison were conducted using the *lsmeans* package [38]. It was revealed that the pitch slope of the homophone (/pɛ1/) of the T1 interrogative SFPs /mɛ1/ was significantly lower (M = -1.42 ERB) than those of the other two conditions. The difference between /mɛ1/ at the end of sentences (M = 2.67 ERB) and its citation form (M = 1.79 ERB) was, however, not significant. Moreover, no significant differences were found across the three conditions for the other three target SFPs (the T1 declarative /lɔ1/, the T3 declarative /a3/, and the T3 interrogative /ma3/) (*ps* > .100). It is therefore suggested that the significant interaction among *condition*, *intonation*, and *tone* was driven by the difference in pitch contour between the T1 interrogative SFPs /mɛ1/ and its homophone /pɛ1/. No other significant interactions were found (*ps* > .100).

3.2. Comparison of the target in Condition 1

Figure 1 shows the mean *f*0 values for the declarative and the interrogative SFPs in Condition 1. The pitch height of each target SFP was first calculated by averaging the nine evenly spaced *f*0 estimates, and was then submitted to a linear mixed effects model. In the linear mixed effects model, the response variable was the pitch heights in ERB-rate. The intercept for *subjects* and *genders* were entered in the random-effects structure. *Intonation* (declarative vs. interrogative) and *tone* (T1 and T3) were entered as fixed effects.

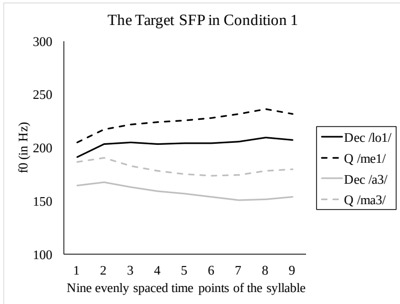

Figure 1: *Mean f0 values of the declarative and the interrogative SFPs.*

A significant effect of *tone* was found, $\chi^2(1) = 167.09$, $p < .001$, where the T1 SFPs bore a higher pitch height than the T3 SFPs. The effect of *intonation* was also found significant, $\chi^2(1) = 28.47$, $p < .001$; higher pitch height was consistently found in interrogative SFPs than the corresponding declarative SFPs. The interaction between *tone* and *intonation* was, however, not significant, $\chi^2(1) = .60$, $p = .439$.

3.3. Global raising

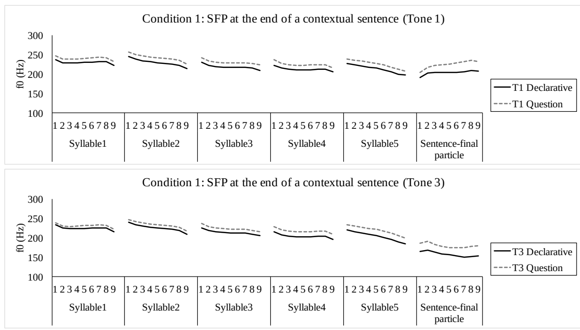

Figure 2: *Mean f0 values of all the syllables in the sentences in Condition 1.*

Figure 2 displays the global *f*0 curves of the sentences in Condition 1 with the SFPs attached to the end, where all six syllables in the sentences are shown in nine evenly spaced time points. Pitch height of each syllable in the carrier sentences was submitted to a linear mixed effects model. *Intonation* (declarative versus interrogative), *tone* (T1 and T3), and *syllable* (the five syllables in the carrier sentence) were entered as fixed effects. The intercept for *subjects* and *genders* were entered in the random-effects structure.

There was a significant effect of *intonation*, $\chi^2(1) = 11.68$, $p < .001$, where pitch levels were higher in the interrogative sentences across all five syllables in the carrier sentences. The effect of *tone* was also significant, $\chi^2(1) = 4.48$, $p < .05$. The pitch values of the carrier sentences bearing the T1 SFPs were consistently higher than those with the T3 SFPs. A significant effect of *syllable* was also found, $\chi^2(4) = 78.54$, $p < .001$, suggesting a global downtrend of pitch over the utterance of the carrier sentences. Post-hoc test with Tukey-adjusted comparison revealed that the pitch heights of the first and second syllables were similar ($p = .947$), and were both higher than that of the third syllable (*ps* < .001). The third syllable was also higher in pitch when compared to the fourth and the fifth syllables (*ps* < .001), while there was no significant difference between the fourth and the fifth syllables ($p = .997$). No other significant interactions were found (*ps* > .100).

4. Discussion

The result of our acoustic analyses showed that the intonation contrasting declarative and interrogative sentences affected the realization of the underlying lexical tones of SFPs. Specifically, interrogative SFPs were consistently produced with higher pitch values than declarative SFPs, despite the fact that they bear the same underlying lexical tone. Moreover, the effect of intonation was not restricted only to the SFPs, but was found in the overall utterance. The global pitch values in questions were higher than the corresponding declarative sentences in Condition 1, where SFPs were attached to the end of sentences.

4.1. The underlying tonal target of SFPs

In the present study, all four target SFPs are shown to bear an underlying tonal target that resembles their homophones when they were both placed in the middle of a sentence (Condition 2 and 3). No significant difference was found between the two in all target SFPs, except for the Tone 1 interrogative SFP /mɛ1/. However, the significant difference found between the citation form of /mɛ1/ and its homophone /pɛ1/ was in fact driven by the impact of the consonantal onset on the pitch of the syllable. The pitch onset tends to become higher when there is a preceding voiceless stop onset consonant. In the present study, the voiceless aspirated bilabial stop /p/ has raised the pitch onset of the homophone /pɛ1/, causing it to bear a falling pitch contour that is different from its counterpart SFP /mɛ1/ with a nasal onset /m/ [39, 40, 41, 42].

Overall, the similarity in pitch slopes within the target syllables across Condition 2 and 3 suggests that SFPs and their homophones have similar underlying pitch target. Such finding conforms to what [25] suggested, where final *f*0 value and final *f*0 velocity were reported to be comparable between the SFPs and their homophones. Our present study compared the pitch slope values of the entire syllables. Despite the methodological difference, both [25] and the present study found that, similar to lexical words, SFPs do bear an inherent lexical tone.

4.2. The effect of intonation on SFPs

We also found that, even when they are attached to the end of sentences, SFPs appear to be similar to their citation form. No significant difference in pitch slope was found between Condition 1 and 2, where the SFPs were placed at the end of a sentence and in the middle of a sentence, respectively. Despite the position at which the SFPs are placed, the pitch contour of SFPs does not bear extra pitch changes and resists the changes imposed by sentence-final intonation. Previous studies reported the effect of intonation on Cantonese lexical tones, where the canonical pitch contour of all lexical tones is modified to bear a rising pitch contour at the final syllable of a question [2, 9, 14, 25]. However, such final rising was absent when the interrogative SFPs were placed at the end of a question. Instead of a final rising effect, the pitch level of the interrogative SFPs is shifted upwards, i.e., they were consistently produced with higher pitch values than declarative SFPs of the same canonical lexical tones.

The local raised pitch values in the interrogative SFPs led to the question of whether it was caused by global raising of questions. Previous studies on Cantonese intonation have proposed that the contrast in intonation signaling questions versus declarative sentences is realized as both a local event and a global event [10, 43, 44]. In addition to the pitch increment within the last syllable, the global pitch curve also shifts upward for questions compared to declarative sentences. In the present study, the global effect of pitch increase in questions was also found. It was revealed that all the syllables consistently bear significantly higher pitch values in questions when compared to the corresponding declarative sentence.

The global raising effect found in the present study therefore implies an interaction between the inherent tone of SFPs and the global pitch increase caused by the question intonation. Specifically, the pitch value realized in SFPs is a combination of the underlying tonal targets and intonation. Different from lexical words where the question intonation is realized as modifying the canonical pitch contour into a rising pitch contour at the end of the syllable (e.g., [6, 9, 14, 15, 16, 17, 18]), the global question intonation caused the underlying tonal target of the SFPs to shift upward. The superimposition of intonation is therefore realized differently on SFPs.

Although SFPs are similar to their homophones in terms of having a basic underlying tonal target, the underlying tonal targets of lexical words and that of SFPs have a different status in the prosodic system of Cantonese. Such a difference could be attributed to the fact that SFPs are not lexical words but grammatical words [11, 17, 20]. The grammatical status of SFPs may have limited the possible application of interrogative final pitch rising. Grammatical words are different from lexical words in terms of phonological properties such that prosodic structure usually makes no reference to grammatical words [45, 46]. The constraint of the grammatical status of words is clearly reflected in SFPs, since the superimposition of intonation is realized differently on SFPs and lexical words. The absence of the final rising in the pitch contour of SFPs therefore suggests that the underlying tonal targets of SFPs and the local final pitch rising effect are in complementary distribution. Instead, the interaction between intonation and the tonal target of SFPs is mainly due to the global pitch raising effect.

5. Conclusions

SFPs are grammatical words, causing them to have a different status than lexical words in the prosodic system. The grammatical status of SFPs have restricted the occurrence of local final pitch rising, which is an essential cue for question signaling in lexical words. It therefore suggests that the underlying tonal targets of SFPs and local pitch changes in questions are in complementary distribution. On the other hand, it was consistently found that questions are produced with a higher global pitch curve than their corresponding declarative sentences. Such upward shift of the pitch value of interrogative SFPs therefore marks an interaction between intonation and lexical tone in the realization of SFPs.

6. Acknowledgement

This research was supported in part by the Early Career Scheme (27402514) and General Research Fund (17673216) from the HKSAR Research Grant Council to Xiuli Tong.

7. References

[1] R. S. Bauer and P. K. Benedict, *Modern Cantonese Phonology*. New York: Mouton de Gruyter, 1997.

[2] Y. Y. Fok-Chan, *A Perceptual Study of Tones in Cantonese*. Hong Kong: Centre of Asian Studies (University of Hong Kong), 1974.

[3] J. Gandour, "Perceptual dimensions of tone: Evidence from Cantonese," *Journal of Chinese Linguistics,* vol. 9, no. 1, pp. 20–36, 1981.

[4] W. Gu, "Tone, intonation, and emphatic stress in L2 Mandarin speech by English and Cantonese learners," in *ICPhS 2015 – 18th International Congress of Phonetic Sciences, August 10–14, Glasgow, UK, Proceedings,* 2015.

[5] Y. Li, T. Lee, and Y. Qian, "Analysis and modeling of F0 contours for Cantonese text-to-speech," *ACM Transactions on Asian Language Information Processing,* vol. 3, no. 3, pp. 169–180, 2004.

[6] A. Fox, K. K. Luke, and O. Nancarrow, "Aspects in intonation in Cantonese," *Journal of Chinese Linguistics,* vol. 36, pp. 321–366, 2008.

[7] D. Hirst, J. C. Wakefield, and H. T. Li, "Does lexical tone restrict the paralinguistic use of pitch? comparing melody metrics for English, French, Mandarin and Cantonese," in *ICPLC 2013 – International Conference on Phonetic of the Languages in China, December 2–4, Hong Kong, Proceedings,* 2013.

[8] W. S. Lee, "The effect of intonation on the citation tones in Cantonese," in TAL 2004 – 1st International Symposium on Tonal Aspects of Languages, March 28–30, Beijing, China, Proceedings, 2004.

[9] J. K.-Y. Ma, V. Ciocca, and T. L. Whitehill, "Effect of intonation on Cantonese lexical tones," *Journal of the Acoustical Society of America,* vol. 120, pp. 3978-3987, 2006.

[10] J. K.-Y. Ma, V. Ciocca, and T. L. Whitehill, "Quantitative analysis of intonation patterns in statements and questions in Cantonese," in *Speech Prosody 2006 – 3rd International Conference on Speech Prosody, May 2–5, Dresden, Germany, Proceedings,* 2006.

[11] S. Matthews and V. Yip, *Cantonese: A Comprehensive Grammar (2nd Ed.)*. London: Routledge, 2011.

[12] M. C. Pennington and N. C. Ellis, "Cantonese speakers' memory for English sentences with prosodic cues," *The Modern Language Journal,* vol. 84, pp. 372–389.

[13] W. L. Wu, "An acoustic phonetic study of the intonation of sentence-final particles in Hong Kong Cantonese," *Asian Social Science,* vol. 4, no. 2, pp. 23–29.

[14] B. R. Xu and P. Mok, "Final rising and global rising in Cantonese intonation," in *ICPhS 2011 – 17th International Congress of Phonetic Sciences, August 17–21, Hong Kong, Proceedings,* 2011, pp. 2173–2176.

[15] Y. R. Chao, "A preliminary study of English intonation (with America variants) and its Chinese equivalent," *The Tsai Yuan Pei Anniversary Volume (Supplement of the Bulletin of the Institute of History and Philosophy,* vol. 1, pp. 105–156, 1932.

[16] P. S. Ding, "From intonation to tone: The case of utterance-final particles 'aa' and 'wo' in Cantonese," *Modern Linguistics,* vol. 1, pp. 36–41, 2013.

[17] H. Kwok, *Sentence particles in Cantonese*. Hong Kong: Centre of Asian Studies (University of Hong Kong), 1984.

[18] Y. Mai, "A brief account of intonation in Cantonese (in Chinese)," in *Research and Teaching in Cantonese (Vol. 3),* D. Cheng and J. Cai, Eds. Guangzhou, China: Zhongsan University, 1998, pp. 32–40.

[19] K.-H. Cheung, The Phonology of Present-day Cantonese (Unpublished PhD thesis). London: University College London, 1986.

[20] X. Fang, Sentence-final Mood Helping-words in Guangzhou Dialect (廣州方言句末語氣助詞). Guangzhou, China: Jinan University Press, 2003.

[21] S.-C. Yau, "Sentential connotations in Cantonese," *Fangyan,* vol. 1, pp. 35–52, 1980.

[22] J. C. Wakefield, The English Equivalents of Cantonese sentence-final particles: A contrastive analysis (Unpublished PhD thesis). Hong Kong: The Hong Kong Polytechnic University, 2010.

[23] J. C. Wakefield, "A floating tone discourse morpheme: The English equivalent of Cantonese *lo1*," *Lingua,* vol. 122, pp. 1739–1762, 2012.

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

[25] W. L. Wu, Cantonese Prosody: Sentence-final Particles and Prosodic Focus (Unpublished PhD thesis). London: University College London, 2013.

[26] L. Zhang, "Segmentless sentence-final particles in Cantonese: An experimental study," *Studies in Chinese Linguistics,* vol. 35, no. 2, pp. 47–60, 2014.

[27] P. Boersma and D. Weenink, *Praat: Doing phonetics by computer (version 6.0.24)*. Amsterdam: University of Amsterdam, 2017.

[28] A. Cruttenden, *Intonation (2nd Ed.)*. Cambridge: Cambridge University Press, 1997.

[29] C. Gussenhoven, *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press, 2004.

[30] D. Hirst and A. D. Cristo, *Intonation Systems: A Survey of Twenty Languages*. Cambridge; Cambridge University Press, 1998.

[31] E. Khouw and V. Ciocca, "Perceptual correlates of Cantonese tones," *Journal of Phonetics,* vol. 35, pp. 104–117, 2007.

[32] R. Wayland and A. Jongman, "Acoustic correlates of breathy and clear vowels: The case of Khmer," *Journal of Phonetics,* vol. 31, no. 2, pp. 181–201, 2003.

[33] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Research,* vol. 47, pp. 103–138, 1990.

[34] D. J. Hermes and J. C. van Gestel, "The frequency scale of speech intonation," *The Journal of Acoustic Society of America,* vol. 90, pp. 97—102, 1991.

[35] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2017.

[36] D. Bates, M. Maechler, B. Bolker, and S. Walker, "lme4: Linear mixed-effects models using Eigen and S4," *R package version,* vol. 1, no. 7, 2015.

[37] J. Fox and S. Weisberg, *An {R} Companion to Applied Regression (2nd Ed.)*. Thousand Oaks: Sage, 2011.

[38] R. V. Lenth, "Least-squares means: The R package lsmeans," *Journal of Statistical Software,* vol. 69, no. 1, pp. 1–33, 2016.

[39] J.-M. Hombert, J. J. Ohala, and W. E. Ewan, "Phonetic explanations for the development of tones," *Language,* vol. 55, pp. 37–58, 1979.

[40] W. A. Lea, "Segmental and suprasegmental influences on fundamental frequency contours," *Consonant Types and Tones,* vol. 1, pp. 15, 1973.

[41] R. N. Ohde, "Fundamental frequency as an acoustic correlate of stop consonant voicing," *The Journal of the Acoustical Society of America,* vol. 75, pp. 224–230, 1984.

[42] N. Umeda, "Influence of segmental factors on fundamental frequency in fluent speech," *The Journal of the Acoustical Society of America,* vol. 70, pp.350–355, 1981.

[43] W. Gu, K. Hirose, and H. Fujisaki, "Analysis of the effects of word emphasis and echo question on F0 contours of Cantonese utterances," in *9th European Conference on Speech Communication and Technology, September 4–6, Lisbon, Portugal, Proceedings, 2005.*

[44] H. Y. D. Lee, The Intonation of Cantonese Declarative Questions (Unpublished MPhil thesis). Cambridge: Cambridge University, 2008.

[45] E. O. Selkirk, Phonology and Syntax: The Relationship between Sound and Structure. Cambridge: The MIT Press, 1986.

[46] E. Selkirk, "The prosodic structure of function words," *Signal to syntax: Bootstrapping from Speech to Grammar in Early Acquisition,* vol. 187, pp. 214, 1996.