



Suprasegmental and prosodic features contributing to perceived accent in heritage Cantonese

Rachel T. Y. Kan¹

¹The Hong Kong Polytechnic University, Hong Kong

rachel.kan@polyu.edu.hk

Abstract

Different suprasegmental and prosodic features impact the perceived native-likeness of speakers, although the exploration of these features have typically focused on second language English. This paper presents a preliminary study investigating the contribution of tonal space and distance between level tones, speech rate, and pause behaviour to the perceived accent of three groups of Cantonese speakers ($n = 5$ each): heritage speakers with high and low native-likeness ratings respectively, and homeland (majority language) speakers. Speech samples were taken from a narrative task, which was also used to obtain the native-likeness ratings from 55 raters. The results showed lower ratings to be associated with smaller tonal space, slower speech rate, and more frequent and longer pauses. Formulae for normalising fundamental frequency (F0) were adjusted and applied to the present data; these changes are also presented and their effectiveness is discussed.

Index Terms: tonal space, speech rate, pauses, accent ratings, Cantonese, heritage language

1. Introduction

Accent ratings have been used as global measurements of the quality or accuracy of production, particularly in second language (L2) learners [1, 2]. These ratings are based on subjective judgments by listeners, who evaluate whether a speaker's speech meets certain expectations. These expectations are usually set with reference to how native speakers speak, and accents are easily associated with status and identity in social contexts [3]. There are various phonetic, phonological, and prosodic features of speech contributing to the perception of a non-/native accent, including overall pitch range [4], word and sentence stress [5, 6], speech rate [7, 8], and articulatory pause duration and frequency [10, 11]. For example, L2 utterances that are either too fast or too slow are perceived as more accented [4, 7, 11], as is speech with frequent and long pauses [4, 9]. Other factors such as speaker characteristics and lexical or grammatical inaccuracy can also play a role [12].

Most studies on this topic target L2 learners of English, so the applicability of the above findings to other target languages as well as the role of other features present in these languages (e.g. tone, pitch accent) are poorly understood. Therefore, the purpose of this study was to examine features contributing to perceived accents in Cantonese. The focus is on the role of lexical tones in foreign accent perception, an area that has not received much research attention. Cantonese has six lexical tones defined by their relative pitch and contours, namely T1 (high level), T2 (high rising), T3 (mid level), T4 (low falling), T5 (low rising), and T6 (low level)

[13]. In neutral, declarative sentences, utterance-level intonation can be observed in the declination in the utterance body (that is, the utterance excluding any sentence final or pause particle) [14]. However, the pitch and contour of tones in the utterance body are generally preserved, even for some question types requiring a rising pitch [15, 16, 17].

Although there are so far no published studies on accentedness in Cantonese speech, there is some data on how the production of Cantonese segments and suprasegmentals is divergent in heritage speakers (HSs) of Cantonese, when compared to majority language speakers in the homeland. HSs do not always show native-like production, due to transfer from the majority language of their host country and/or incomplete acquisition of the heritage language [18]. HSs of Cantonese show reduced tonal space and merging between tones [19, 20], which could lead to perceived accents. At the same time, merging of T2-5 and T3-6 has been observed in adult homeland speakers of different varieties of Cantonese, so studies on their production of tones are also needed [21, 22]. Bilingual Cantonese speakers from Hong Kong also show more errors in vowel production compared to predominantly monolingual speakers, such as in smaller length and vowel contrasts [23, 24]; vowel length is not contrastive in Cantonese, but some long-short vowel pairs occur in complementary distribution depending on the coda.

Speakers with phonetic or phonological production that is qualitatively different from native speakers or any accepted or standard varieties will be perceived as having a foreign accent. In fact, segmental/suprasegmental features predict accent ratings even more strongly than lexical or grammatical divergences [25, 26]. Therefore, it is predicted that Cantonese features where such qualitative differences were previously observed (e.g. tones, vowel length) can be linked to differences in accent ratings, and that features identified previously for L2 English and that are present in Cantonese (e.g. speech rate, pause duration, frequency) would also apply.

Therefore, this paper joined findings from more general literature to research on Cantonese-specific features, to investigate the relation between the following features and accent ratings in Cantonese: (1) tonal space and distance, (2) speech rate, and (3) pause duration and frequency. HSs of Cantonese who were perceived as more native-like and less native-like respectively were compared to majority language speakers in the homeland. Based on previous literature, the HSs perceived as less native-like would show reduced tonal space, slower speech, and more frequent and longer pauses compared to speakers perceived as more native-like, due to stronger transfer from English and lower Cantonese proficiency.

Accent ratings are typically obtained using samples of spontaneous speech, while the production of tones is typically evaluated based on tokens produced in controlled carrier sentences or in isolation. Since it was necessary to use spontaneous speech here, this paper also tested whether current methods for processing and analysing tones, which are used on tokens produced in more controlled environments, could also be applied to spontaneous speech.

2. Method

2.1. Participants

Participants were heritage speakers (HSs) of Cantonese living in the US, who were selected for the present paper based on native-like ratings obtained in a separate study. These ratings were given by adult native speakers of Cantonese in Hong Kong ($n = 30$) and Guangzhou ($n = 22$), representing the two main varieties of Cantonese.

Data on the HSs' linguistic and family background was collected through a language background questionnaire to parents adapted from the BiLingual Language Experience Calculator [27]. The first group of HSs were five children who received the highest native-likeness ratings ("HI" group; mean rating 4.9 on a scale of 1-6, with 6 as the most native-like). They were all female, aged 7;1-10;11 ($M = 9;1$), all born in the US. The second group of HSs were five children who received the lowest native-likeness ratings ("LO" group; mean native-likeness rating = 2.8). These included 3 male and 2 female speakers, aged 5;11-10;7 ($M = 8;9$). Three were born in the US, one in Hong Kong (age of arrival = 7;11; length of residence = 2.3 years), and one in Mainland China (age of arrival = 1;6; length of residence = 8.3 years). Both groups heard and spoke predominantly Cantonese at home (heard Cantonese: HI: 81% of the time; LO: 81%; spoke Cantonese: HI: 73%; LO: 74%), but attended local primary schools where English was used with teachers and other students.

The control group ("CON" group) was selected to best match the HI group in both native-likeness ratings ($M = 5.6$) and age (6;5-11;4; $M = 8;11$). They were all children born and raised in Hong Kong (2 female 2 male). They used mainly Cantonese at home (overall 80% of the time) and in school (69% of the time).

The number of (grammatical) mistakes was not found to be a significant predictor of native-likeness ratings and grammatical features are not investigated here, so it was not considered in the selection of participants.

2.2. Speech data

The speech data came from a narrative task eliciting Cantonese production using the 'Frog, Where Are You?' story book [28], which only has pictures and no text. Participants were instructed to describe what was happening in the pictures in Cantonese. Samples were taken to obtain the nativelikeness ratings. For the present paper, the whole recording for each child was used, but utterances with rising intonation were excluded. Data collection took place in a quiet classroom in the children's schools, and recordings were made using a desktop microphone connected to a digital recorder.

2.3. Analysis

2.3.1. Tonal space and distance

The pitch level of a tone is measured by its fundamental frequency (F0). Tonal space was measured by the pitch range between the average F0 of the highest and lowest tones (T1 and T4 respectively) for each speaker (adapted from [20]). The spatial relationship between the three level tones was measured by taking the F0 difference between T1-T3, T1-T6, and T3-T6. Tokens matching the target tones were extracted from the narrative recordings. F0 values were obtained by marking the sonorous parts of target syllables manually in Praat [29] and extracting the average F0 for each token using ProsodyPro [30]. Tokens that were creaky, too short, or from exaggerated utterances (leading to dramatic change in pitch levels) were excluded. Tokens with unreliable F0 estimates (e.g. due to background noise or unstable values across a syllable) were also excluded. The number of tokens valid for analysis for each group were CON: 503, HI: 1539, LO: 1544.

Speakers' tonal patterns can be more easily understood and compared by using Chao's tone letters [13], which is a tone notation/transcription scale that divides a speaker's pitch range (tonal range) into 4 equal bands, and marks the boundaries with 1 (lowest) to 5 (highest). Although these intervals and the distance between them are not matched precisely to speech data, the Chao system is still widely-used to describe tones, including by the International Phonetic Association. The numerical equivalents of the tone letters are used in this paper, and were obtained using the following formulae for Cantonese [20]:

$$N = 39.86 * (\log f0_i - \log f0_{min}) \quad (1)$$

$$\text{Tone letter scale} = \frac{N}{2} + 1 \quad (2)$$

where $F0_i$ is the F0 estimate of the target token, $F0_{min}$ is the minimum (lowest) F0 estimate obtained for a particular speaker, and the interval between each point on the scale as 2.5 semitones (following [20]). However, the resulting tone letters often exceeded 5, which is the highest point on Chao's scale. To illustrate, Figure 1 shows the tokens produced by a CON group participant in different tones. Here and in the following, only the level tones will be shown. The corresponding Chao tone letters are '55' (T1 high level), '33' (T3 mid level), and '22' (T6 low level). (The high falling/'51' variant of T1 has not been widely used in Hong Kong Cantonese since back in the nineties [31, 32].)

A possible reason is that a speaker's pitch range (and therefore the appropriate F0 values for different tones) changes across time during spontaneous speech, so that a single $F0_{min}$ is not appropriate for recordings of narratives. Therefore, Formula 1 was adjusted so that the $F0_{min}$ applied to a particular token is the minimum value observed in the utterance containing that token, although this method could only be used for utterances containing T4, the lowest tone. The potential of local normalisation was explored, and the analysis also used raw F0 values.

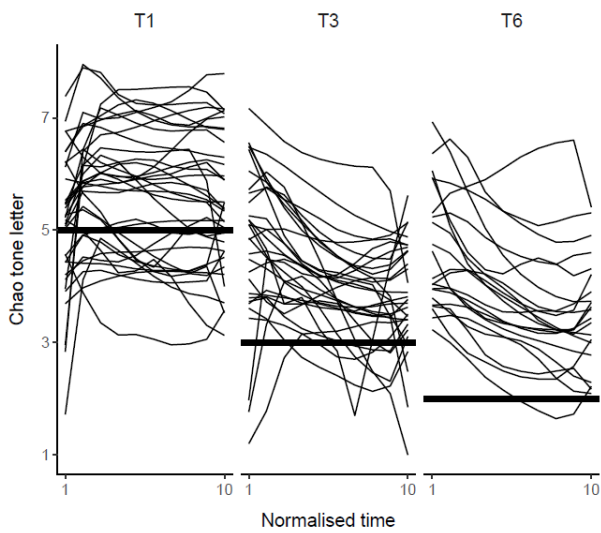


Figure 1: Tone contours of tokens produced by a CON group participant, represented using Chao tone letters. Each token was sampled at 10 equidistant intervals for presentation. The horizontal lines represent the conventional tone contours [13].

2.3.2. Speech rate

Speech rate was calculated by dividing the number of syllables produced by the duration of the utterance, both excluding and including pauses [4, 5]. The number of syllables was counted including English words, which were present in the recordings of three participants in the HI group (2-8 syllables) and of three participants in the LO group (19-26 syllables).

2.3.3. Pausing behaviour

The absolute frequency of articulatory pauses was calculated by counting the number of pauses (breaks between words of <0.1s including filled pauses) within utterances [33]. The relative frequency of pauses compared to the number of syllables was also included, as participants may not produce the same length (in words and duration) of narrative for various reasons. The relative duration of pauses was obtained by dividing the combined length of all pauses by the combined length of all utterances (cf. [4]).

3. Results

3.1. Local normalisation to obtain Chao tone letters

Only tokens in utterances with at least one T4 word could be included when using local normalisation to obtain Chao tone letters, so the data set for each group was greatly reduced (Table 1). Two CON group participants were excluded as T6 tokens were no longer present in the data. Although T4 was the most common tone of the tokens used as $F0_{\min}$, tokens in T2 and T6 were used if they had a lower (raw) $F0$ than the T4 token(s) in the same utterance.

Figure 2 illustrates the Chao tone letters obtained using local normalisation for the same CON participant in Figure 1. Table 2 shows that the mean Chao tone letter values for different tones did not match well with the conventional values (the two numbers represent the initial and final values of each tone; these three are level tones so have the same

initial and final values [13]). The order of tones (i.e. relative height) was preserved ($T1 > T3 > T6$) for all three groups, but the intervals between tones diverged. Some values for T1 exceeded 5, the highest letter on the scale. The SDs were also large relative to conventional interval between tones (e.g. the distance between T3 and T6 is one interval).

Table 1: Size of dataset to which local normalisation could be applied

	No. utterances	No. tokens	Difference in no. tokens
CON	14	102	-79%
HI	79	929	-40%
LO	76	652	-58%

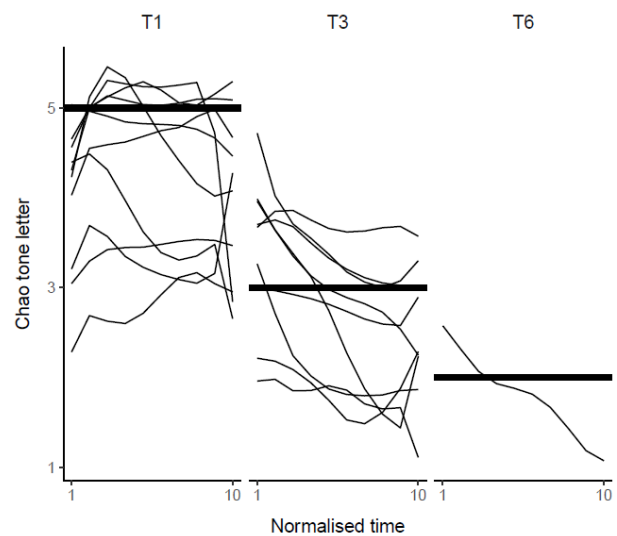


Figure 2: Tone contours of tokens produced by a CON group participant, represented using Chao tone letters. The horizontal lines represent the conventional tone contours [13].

Table 2: Mean Chao tone letter values of tokens for different tones compared to the convention (SD)

	T1	T3	T6
Convention	55	33	22
CON	4.5 (0.57)	2.1 (0.52)	1.62 (0.27)
HI	5.12 (0.54)	3.29 (0.55)	2.44 (0.69)
LO	4.77 (1.15)	3.15 (0.59)	2.86 (1.1)

3.2. Tonal space and distance between level tones

Tables 3 and 4 show the tonal space and distance between level tones produced by the three groups, calculated using Chao tone letters and $F0$ values respectively. Values from the literature are also included for reference: convention from [13] in Table 3 and adult native male and female speakers in [20] in Table 4. (The tone letters for T4 are '21'.) The results from the two calculation methods do not show the same pattern. Based on Chao tone letters, the HI group has a largest tonal space, followed by CON and then LO. A similar pattern is observed for the distance between T3-T6. Between T1-T3 and between T1-T6, the distance is largest for the CON group,

followed by HI, then LO. Based on F0 values, there is little difference between groups.

Table 3: *Measurements of tonal space and distance calculated using Chao tone letters (SD)*

	T1-T4	T1-T3	T1-T6	T3-T6
Reference	3 or 4	2	3	1
CON	2.94 (0.27)	2.40 (1.09)	2.88 (0.29)	0.47 (0.79)
HI	3.18 (0.80)	1.83 (0.44)	2.67 (0.41)	0.84 (0.27)
LO	2.43 (1.59)	1.62 (1.21)	1.90 (1.21)	0.28 (0.94)

Table 4: *Measurements of tonal space and distance calculated using F0 values (SD)*

	T1-T4	T1-T3	T1-T6	T3-T6
Reference				
(Female)	94.77	52.7	75.2	22.5
(Male)	52.44	27.03	42.88	15.85
CON	35.61 (10.71)	31.83 (9.65)	35.61 (10.71)	3.78 (1.07)
HI	37.10 (8.79)	29.12 (10.74)	37.10 (8.79)	7.98 (2.79)
LO	38.44 (16.33)	33.33 (11.79)	38.44 (16.33)	5.11 (9.11)

Finally, mixed effects regression analyses were conducted to test the differentiation between tones and whether this differentiation was the same across groups. A linear mixed model was fitted using the lme4 package [34] in R [35], with the Chao tone letters as the dependent variable. Participant was entered as a random factor. Fixed factors were Tone, dummy coded as T1 (reference level), T3, and T6, and Group, dummy coded as CON (reference level), HI, and LO. An interaction between Tone and Group was also included.

The results showed that for CON, T3 and T6 both differed significantly from T1 [$B = -2.49$, $t(800.33) = -5.8$, $p < 0.001$; $B = -2.90$, $t(800.88) = -5.5$, $p < 0.001$]. Post hoc Tukey pairwise comparisons (calculated using the multcomp package, [36]) indicated similar results for both LO and HI. The difference between T3-T6 was significant for HI ($p < 0.001$), but not for CON and LO ($ps > 0.05$). All other comparisons and the interaction terms returned no significant effects ($ps > .05$).

3.3. Speech rate

The speech rate of the CON and HI groups were more similar (with pauses: 1.89-2.68 ($M = 2.35$) and 1.58-2.96 ($M = 2.27$) syllables/second respectively; without pauses: 2.89-3.27 ($M = 2.98$) and 2.70-3.77 ($M = 3.26$) syllables/second respectively), while the speech rate of the LO group was relatively low (with pauses: 1.27-1.63 ($M = 1.38$) syllables/second; without pauses: 2.11-2.80 ($M = 2.41$) syllables/second).

3.4. Pausing behaviour

The CON group paused the least frequently (27-35 pauses, $M = 29.8$), followed by the L group: 43-119, $M = 84.6$), followed by HI group (58-187, $M = 94.8$). However, the HI group included one participant who paused 187 times. Excluding this outlier, the HI group paused 58-92 times ($M = 71.75$), making them similar to the LO group.

The CON group paused less often per syllable produced (0.10-0.15s per syllable, $M = 0.12s$), while the two HS groups were more similar (HI: 0.19-0.47s, $M = 0.28s$; LO: 0.20-0.45s, $M = 0.32s$). Considering the relative duration of pauses (length of pauses divided by length of utterances), the CON group's pauses were the shortest (0.16-0.49s, $M = 0.27s$), followed by the HI group (0.27-0.50s, $M = 0.46s$), whereas the LO group paused for the longest (0.55-1.16s, $M = 0.76s$) relative to the length of utterances they produced.

4. General Discussion

This paper investigated the role of features in perceived accents in Cantonese, and compared HSs of Cantonese with high and low native-likeness ratings to majority language speaker peers. Data was taken from narrative recordings which had also been used to obtain the native-likeness ratings. The results showed that when using Chao tone values, the tonal space and distance between level tones was smaller in the HSs with low native-likeness scores (LO group), indicating that reduced tonal space could be linked to more accented speech [19, 20]. The difference between the CON and HI groups is not clear, although the difference between their native-likeness scores (0.5 on a scale of 1-6) is also smaller compared to the difference between the HI and LO groups (2.1). These initial findings suggest a role for tone in perceived accents, although they are limited by possible unreliability in the F0 estimations.

This paper also explored an adjusted method of obtaining Chao tone letters, by applying a local value during normalisation. While it appears to be an improvement over using a single value across a recording, various issues remain. For example, the Chao tone letters did not meet theoretical expectations completely (e.g. with T1 tokens exceeding the highest tone letter of 5). There was also considerable variation even within one participant. One way of further refining the data is to control the position of each token within an utterance, and hence account for declination. (This was not possible in this paper given the spontaneous nature of the data and limited sample.) These limitations reflect the difficulty in analysing spontaneous speech recorded in a relatively naturalistic environment. The use of pitch for pragmatic or dramatic effects could be quite common in story-telling.

For both speech rate and pausing behaviour, findings were consistent with previous literature in that more accented speech (LO group) tended to be slower and contain more and longer pauses, at least compared to the CON group [7, 9, 11]. The HI group was expected to rank middle for measurements that affect perceived accent, so contrasts with other groups were less clear. Their status can be clarified by increasing the sample size for all groups, and further incorporating statistical analyses. One point to note is that the native-likeness ratings had been obtained based on 10s samples and pauses of >1s had been reduced to 1s, due to considerations for consistency across samples and time. While the present preliminary results are based on a limited sample, they suggest that previously identified features affecting accent ratings apply to Cantonese, and are promising regarding further, full-fledged studies.

5. Acknowledgements

This work is partially funded by ESSEXLab and the Hong Kong Polytechnic University Postdoctoral Fellowships Scheme. Comments and suggestions from the reviewers are gratefully acknowledged.

6. References

- [1] M. J. Munro and T. M. Derwing, "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners," *Language Learning*, vol. 45, no. 1, pp. 73–97, 1995.
- [2] P. Trofimovich and W. Baker, "Learning prosody and fluency characteristics of L2 speech: The effect of experience on child learners' acquisition of five suprasegmentals," *Applied Psycholinguistics*, vol. 28, pp. 251–276, 2007.
- [3] T. M. Derwing and M. J. Munro, "Putting accent in its place: Rethinking obstacles to communication," *Language Teaching*, vol. 42, no. 4, pp. 476–490, 2009.
- [4] O. Kang, "Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness," *System*, vol. 38, no. 2, pp. 301–315, 2010.
- [5] J. Field, "Intelligibility and the listener: The role of lexical stress," *TESOL Quarterly*, vol. 39, no. 3, pp. 399–423, 2005.
- [6] L. D. Hahn, "Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals," *TESOL Quarterly*, vol. 38, no. 2, pp. 201–223, 2004.
- [7] M. J. Munro and T. M. Derwing, "The Effects of Speaking Rate on Listener Evaluations of Native and Foreign-Accented Speech," *Language Learning*, vol. 48, no. 2, pp. 159–182, 1998.
- [8] M. J. Munro and T. M. Derwing, "Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate," *Studies in Second Language Acquisition*, vol. 23, pp. 451–468, 2001.
- [9] X. Liu and J.-K. Lee, "The Contribution of Prosody to the Foreign Accent of Chinese Talkers' English Speech," *Journal of the Korean Society of Speech Sciences*, vol. 4, no. 3, pp. 59–73, 2012.
- [10] M. J. Munro and T. M. Derwing, "Evaluations of foreign accent in extemporaneous and read material," *Language Testing*, vol. 11, no. 3, pp. 253–266, 1994.
- [11] P. Trofimovich and W. Baker, "Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech," *Studies in Second Language Acquisition*, vol. 28, pp. 1–30, 2006.
- [12] M. S. Schmid and H. Hopp, "Comparing foreign accent in L1 attrition and L2 acquisition: Range and rater effects," *Language Testing*, vol. 31, no. 3, pp. 367–388, 2014.
- [13] Y. R. Chao, *Cantonese primer*. Cambridge, M.A.: Harvard University Press, 1947.
- [14] A. Fox, K. K. Luke, and O. Nancarrow, "Aspects of intonation in Cantonese," *Journal of Chinese Linguistics*, vol. 36, pp. 321–367, 2008.
- [15] Chow, S. *A phonetic study of the intonation in Hong Kong Cantonese*. Unpublished MPhil thesis, City University of Hong Kong, 2002.
- [16] 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, no. 6, pp. 3978–3987, 2006.
- [17] S. Matthews and V. Yip, *Cantonese: A comprehensive grammar*, 2nd ed. London: Routledge, 2001.
- [18] S. A. Montrul, *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins, 2008.
- [19] M. H. B. Leung and H. Goad, "The production of the low-level and low-falling tone merger by heritage Cantonese speakers in Toronto." In prep.
- [20] K. L. C. So, "Tonal production and perception of Canadian raised Cantonese speakers (Unpublished MA dissertation)," Simon Fraser University, Burnaby, 2000.
- [21] R. S. Y. Fung and C. K. C. Lee, "Tone mergers in Hong Kong Cantonese: An asymmetry of production and perception", *The Journal of the Acoustical Society of America*, vol. 146, EL424, 2019.
- [22] J. Ou. Tone merger in Guangzhou Cantonese. Unpublished MPhil thesis, Hong Kong Polytechnic University, 2012.
- [23] A. Chan, W. L. Wu, J. Yeung, and C. Choi, "Cantonese vowel production in pre-school Urdu-Cantonese bilingual minority children (poster)," *Conference on Speech, Language and Communication Science and Disorders across the Life Span 2017, Aug 10, Hong Kong*, 2017.
- [24] P. Mok, C. W. Leung, C. Lan, and A. C. L. Yu, "The acquisition of Cantonese vowel length contrast and vowel rounding contrast by South Asian students in Hong Kong," *19th International Congress of Phonetic Sciences (ICPhS 2019), August 5-9, Melbourne, Australia, Proceedings*, 2019, pp. 1–5.
- [25] K. Saito, P. Trofimovich, and T. Isaacs, "Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels," *Applied Psycholinguistics*, vol. 37, no. 2, pp. 217–240, 2016.
- [26] R. Wayland, "Non-native production of Thai: Acoustic measurements and accentedness ratings," *Applied Linguistics*, vol. 18, no. 3, pp. 345–373, 1997.
- [27] S. Unsworth, "Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender," *Bilingualism: Language and Cognition*, vol. 16, pp. 86–110, 2013.
- [28] M. Mayer, *Frog, where are you?* New York: Dial Press, 1969.
- [29] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer." 2019.
- [30] Y. Xu, "ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis," *Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), August 30, Aix-en-Provence, France, Proceedings*, 2013, pp. 7–10.
- [31] R. S. Bauer and P. K. Benedict, *Modern Cantonese phonology*. Berlin, New York: Mouton de Gruyter, 1997.
- [32] L. K. H. So, "Tonal changes in Hong Kong Cantonese". *Current Issues in Language and Society*, vol. 3, no. 2, 186–189, 1996.
- [33] Riazantseva, "Second language proficiency and pausing: A study of Russian speakers of English," *Studies in Second Language and Acquisition*, vol. 23, pp. 497–526, 2001.
- [34] D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.
- [35] R Core Team, "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, 2017.
- [36] T. Hothorn, F. Bretz, and P. Westfall, "Simultaneous inference in general parametric models," *Biometrical Journal*, vol. 50, no. 3, pp. 346–363, 2008.