Changing against tone merging trends in community? The case of C. Y. Leung

Ziqi Chen
The Hong Kong Polytechnic University/
Hung Hom, Hong Kong
`scarlett.chen@connect.polyu.hk`

Yao Yao
The Hong Kong Polytechnic University/
Hung Hom, Hong Kong
`ctyao.yao@polyu.edu.hk`

Alan C. L. Yu
University of Chicago/
Chicago, IL, US
`aclyu@uchicago.edu`

Abstract

Compared to the voluminous literature of sound change and variation across speakers and generations, much less is understood about change and variation within individuals across the lifespan. In this paper, we present a study of the Cantonese speech of a public figure, Mr. Chun-ying Leung (1954-), recorded over a span of three decades (1984-2013), which coincides with a period when multiple sound changes were in progress in Hong Kong Cantonese, including the merging of tones. Assisted with automatic segmentation and annotation, 13,389 tokens representing 878 unique characters with unchecked tones were measured for F0. Our results show that instead of merging tones, Leung made increasingly greater distinction in the past three decades between potential tone-merging targets, exhibiting sound changes in the opposite directions of the ongoing changes in the community.

1   Introduction

Compared to the voluminous literature of sound change and variation across speakers and generations, much less is understood about change and variation within individuals across the lifespan. A few panel studies (e.g. Sankoff & Blondeau, 2007) followed groups of adult speakers for decades, and found that while the majority of the speakers show stability of pronunciation over the years, a minority of speakers do change, especially when the sound change takes on social meanings. Another line of studies tracked the
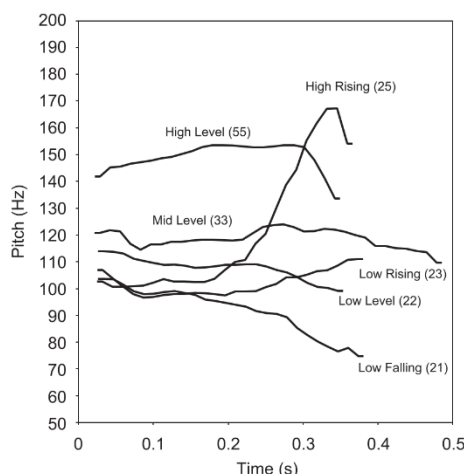

Figure 1. F0 contours for HKC tones produced on the syllable [ji] by a male (Francis, Ciocca, Ma, & Fenn, 2008)

speech of eminent individuals over a long stretch of time. Harrington and colleagues (Harrington, Palethorpe & Watson, 2000) showed significant changes in Queen Elizabeth II's vowels across her lifespan, in a direction consistent with the changes in progress in the broader language community. Similar studies of Ruth Bader Ginsburg (Shapp, LaFave, & Singler, 2014) and Noam Chomsky (Kwon, 2018) also found significant changes across the lifespan.

Past apparent time studies reported Cantonese sound changes in progress, tone merging as one type of them (Mok, Zuo, & Wong, 2013; To, Mcleod, & Cheung, 2015). The studies suggested the merging of the tones in HKC: originally 6 tones, merging observed between low rising T2 and high rising T5, and among mid-level T3, low-

level T6, and low-falling T4 (Mok et al., 2013). The studies used synchronic variations to represent diachronic changes, based on the apparent time hypothesis that one's language acquired in childhood would remain stable in adulthood (Labov, 1999, p. 45). It was also assumed that the speech of an adult should record the communal trends of the period when he was a child; and thus, by collecting data across age groups, researchers could compare linguistic patterns in different periods of a community. Given the significant changes across the lifespan reported by the above-mentioned real time studies, the validity of the hypothesis and the generality of language variations revealed by "synchronic approach" were challenged (Bailey, 2002). As a result, more real time evidence is needed in our understanding of Cantonese sound changes.

The current work presents a study of the Cantonese speech of Mr. Chun-ying Leung (1954-), a former Chief Executive of Hong Kong (2012-2017), recorded over a span of three decades (1984-2013). The window of investigation focuses on his tonal production patterns spanning thirty years, which coincides with a time period when multiple sound changes, including tone merging, were in progress in Hong Kong Cantonese (HKC). The study is motivated by two research questions: (1) Did the adult change his accent over time? (2) If so, did he change his accent in the same direction as the communal trends? In this paper, Section 2 introduces the methods of data collection and processing. Section 3 shows the changes of the tonal curves over three decades. Section 3 explores possible explanations for the changes. Section 4 concludes this paper.

2 Methods

2.1 Subject selection

Two criteria were set in subject selection. The selected individual should be born and raised in Hong Kong and speak HKC without an obvious foreign accent. Besides, the speech recordings of the individual should be available and overall present adequate time depth.

Based on the criteria, Mr. Chun-ying Leung (1954-, hereafter abbreviated as LCY), a former Chief Executive of Hong Kong (2012-2017), was selected in the study. No obvious accent out of

Hong Kong was reported for LCY. Four native HKC speakers were invited to perform subjective judgement of the accent of LCY. Half of them were born in 1990s, the other in 1960s. After listening two 3-minute long recordings by LCY in 1988 and 2013, the participants were asked to answer intuitively whether the individual in the recordings spoke native HKC or a foreign accent. All the four participants reported that the person spoke native HKC. Besides, LCY' recordings are available online or in library, and the recorded time in the data collected spans from 1984 to 2013, favoring the current research with adequate time depth of 30 years almost.

2.2 Recording collection

A total of 11 video recordings of LCY's Cantonese speech, given as interviews or public speeches, were collected from the public portal YouTube or DVD resources in library. The video recordings were converted into audio files, with mono channel, 16,000 sampling rates, and then saved in 32-bit WAV format. The recordings were further edited to exclude the speech by the interlocutors of LCY and the overlapping speech in turn-taking.

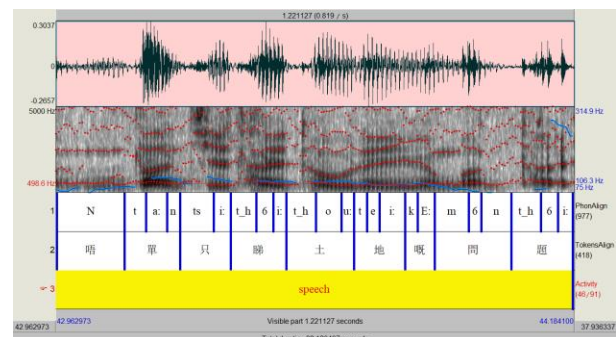### 2.4 Automatic alignment

Figure 2. Automatic annotation results

Using the software tool SPPAS (Bigi & Hirst, 2012), the speech recordings were automatically segmented at phoneme and syllable levels (see Figure 2). They were also annotated through aligning provided transcription texts to segments accordingly. The recordings were transcribed verbatim in traditional Chinese. For nonstandard characters in spoken HKC, the study referred to the Cantonese lexicon equipped in SPPAS (Fung & Bigi, 2015).

2.5 Further annotation

Since automatic annotation was realized at segmental level only, we further annotated citation tones based on a Chinese character database for Cantonese (粵語審音配詞字庫) (Kwan, 2003). Based on the word list of each tone category, a citation tone label was assigned to each token.

For the homographs (in this study, only the characters with more than one possible tone), we manually annotated the tokens occurring in the 1980s, because the size of data in 1980s was limited. Referring to the database (Kwan, 2003), syntactic functions, syllabic pronunciations, and word meanings were used as the disambiguation criteria for each character. In the annotation results of the 1980s, if there were more than one token of a character labelled as the same tone, the tone was deemed as the only one for the character, and the rules were applied to the whole data set. The rest of homographic tokens unaffected by the manual correction and rule propagation was labelled as Null and excluded from following analysis.

2.6 F0 measurement

The automatic segmentation was followed by adjustments to only mark the beginning and end of periodicity of each token. As a result, for each syllabic token, the segmentation started from the onset time of its sonorant (if its initial consonant is sonorant) or vowel (if its initial consonant is obstruent) and ended up with the stop time of its vowel.

Ten measurement points for F0 values were taken between the two boundaries at equal distance, by using the script ProsodyPro (Xu, 2013) on Praat (Boersma & Weenink, 2018). It resulted in the time-normalized data for each tone. In avoidance of the F0 measurement interfered by initial consonants and creakiness: the data from the first and the last measurement points were excluded, and thus only data from eight measurement points were used for analysis.

3 Results

The whole dataset originally included 16,815 tokens produced by LCY, whereas the study only focused on unchecked tones, represented by 15,352 tokens in total. For the data distribution of unchecked tone tokens, most of the data points fell

into the normal pitch range for male, around 100 to 150 Hz for F0 means (Traunmüller & Eriksson, 1995).

The original dataset included extremely high F0 values, larger than 3000 Hz. Given the fixed frequency range of the voice fundamental in an individual' speech, instead of random distribution, the data points falling outside of the normal range were attributed to measurement errors. They could be caused by creakiness during phonation or imperfect inter-pausal segmentation, which included white noise residuals at the beginning or ending portions of a token. Meanwhile, the minimum values were overall larger than 60 Hz, which were within a male's normal pitch range. Thus, the original data distribution was right skewed: the mean value across all measurement points was around 119 Hz, 23 Hz larger than the median, 96 Hz.
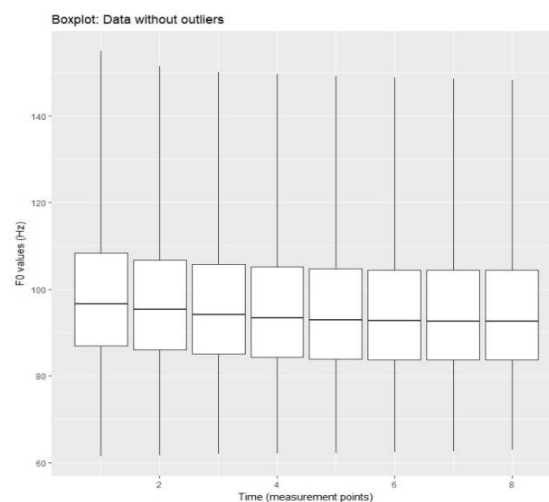
3.1 Outlier identification

Figure 3. Outlier-removed data distribution

To prevent the analysis from being biased by measurement errors, the tokens with extreme values were excluded. According to the "1.5 × IQR" rule for outlier identification (Hubert & Van der Veeken, 2008), the study calculated the first and third quartiles (i.e. Q1 and Q3) at each measurement points, and their inter-quartile range (IQR) as the difference between Q1 and Q3. Then "Q1 - (1.5 × IQR)" and "Q3 + (1.5 × IQR)" were set as the lower and the upper boundaries for data at each measurement point, and the data points outside of the boundaries were identified as

outliers. For each token, among its eight measurement points, as long as it included one data point identified as an outlier, the token was excluded from following analysis.

13,389 out of 15,352 unchecked tone tokens remained after removing outliers, resulting in the loss of 12.7% of the data. Figure 3 illustrates the data distribution after removing outliers from raw data. Consistent with the normal pitch range for males (Traunmüller & Eriksson, 1995), the F0 values ranged from approximately 60 to 150 Hz across the eight measurement points, showing almost symmetric distribution, with the mean values around 96 Hz, close to the median values around 93 Hz.

3.2 Differences between year groups

The pitch levels increased from the 1980s to the 1990s and after 2000. On average, the pitch level in 1980s placed around 87.5 Hz was increased to 92.5 Hz in the 1990s, and 100 Hz after 2000. Apart from the increasing pitch levels, the pitch ranges expanded over time: 15.12, 17.18 and 25.38 Hz in the 1980s, the 1990s and after 2000 respectively (the difference between the maximum and the minimum means in each year group), shown in Figure 4.

The study also investigated whether the expansion of pitch range and the growth of pitch level were related to variation in delivery style, which was indexed by speech rates. It calculated the speech rates (wpm, words per minute) in each recording, by selecting a segment of continuous speech locating in the middle of a recording and counting the number of tokens produced within the segment. Then the speech rate on weighted average for each year group was obtained from the summation of the year speech rates weighted by their duration percentage in the year group, e.g. the speech rate in 1980s: $259.21 \times 0.12 + 223.47 \times 0.79 + 249.14 \times 0.09 = 230.12$ (wpm).

Year	Speech rate (wpm)	Pitch range (Hz)	Pitch level (Hz)
1980s	230.12	15.12	87.27
1990s	274.62	17.17	91.88
After2000	273.79	25.37	99.85

Table 1: Speech rates

The changing patterns of speech rates were inconsistent with those of pitch ranges and levels.

As shown in Table 1, from the 1980s to the 1990s, the speech rates significantly rose from 230.12 to 274.62 wpm whereas the increases in pitch range and level were merely 2.05 and 4.61 Hz. More importantly, the speech rate slightly dropped to 273.79 after 2000, whereas the significant surges were witnessed for both pitch range and level by 8.2 and 7.67 Hz respectively. Thus, inferred from the inconsistency of changing patterns, there is no correlation between the speech rates and the pitch ranges and levels.
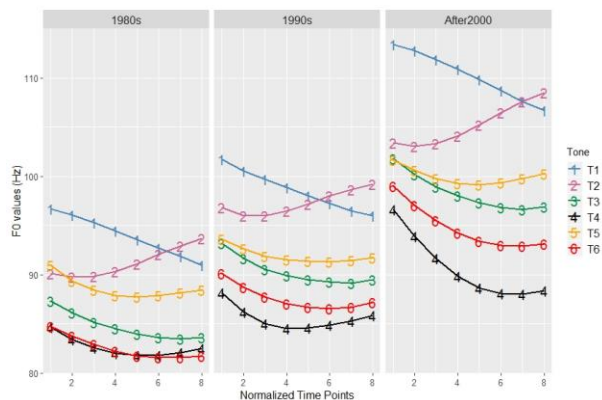

Figure 4. Tone contours of year groups

3.3 T2 and T5

Canonically, T2 and T5 should start with near onset values, with one reaching the highest value within one's normal pitch range and one achieving a slight low-rising (Bauer & Benedict, 1997). Whereas in the speech production of LCY, generally, the gap between the tonal contrast were widened over time, especially in the final portion of time points.
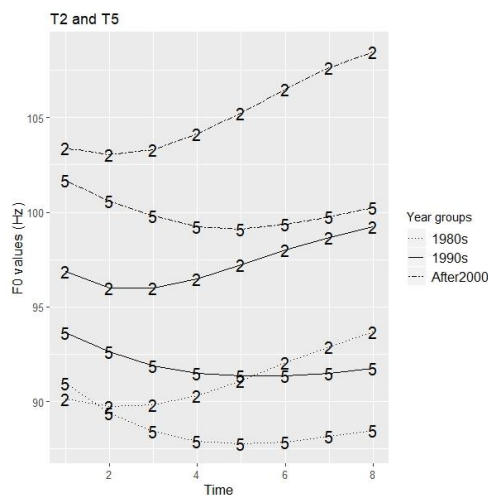

Figure 5: T2 and T5

The tonal contrast T2 and T5 became more and more distinct over time, as shown from Figure 5. In 1980s, the tone pair started with similar onset values with T2 0.82 Hz less than T5 slightly, and then the curves crossed over between the points 1 and 2 with the distinction continuously increased to 5.20 Hz at the end. The difference in 1990s was widened by 4.26 Hz, from 3.21 to 7.47 Hz. After 2000 the pair was further separated. Though started with adjacent onsets with merely 1.72 Hz difference, T2 surged to 108.48 Hz at the end, showing an evident elevation, whereas T5 only demonstrated a modest increase from 99.12 to 100.23 Hz in the last half portion of time points.

3.4 T4 and T6

Both T4 and T6, by convention, are low in pitch height with T6 minimally higher than T4 at the beginning; but they largely differ in pitch contour: T4 shows a decline (low-falling), almost reaching to the bottom of one's pitch range, while T6 remains low-level (Bauer & Benedict, 1997) .
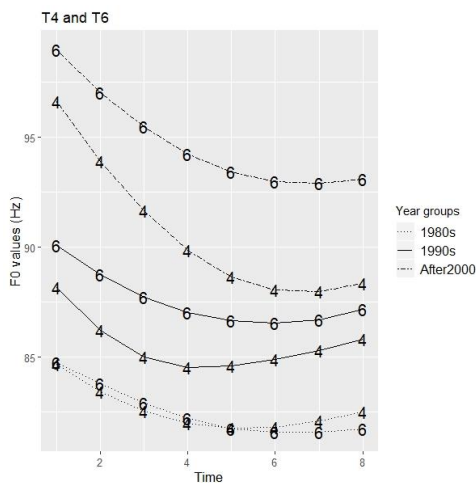

Figure 6. T4 and T6

The tonal contrast T4 and T6 showed the most evident process of differentiation over time, starting with a merging state and ending up with clear separation. Both T4 and T6 in the 1980s were placed at the bottom of his pitch range at around 82.5 Hz, and they nearly overlapped and thus merged together, shown as the dotted curves in Figure 6. In the 1990s, the pair preserved approximately the 2 Hz difference in pitch height but parallel pitch contours: T4 decreased from 88.16 to 84.52 Hz from the points 1 to 3 but showed a modest increase to 85.83 Hz between in

the remaining part; T6 also declined from 90 Hz in its initial portion, but increase a bit to remain level at around 87 Hz between the points 3 and 8. The tonal pair after 2000 was clearly separated in pitch height, at the distance of around 5 Hz in the final portion. Compared to T4 noticeably dropping from 96.62 to 88.36 Hz, the pitch range for T6 was relatively narrowed, from 98.99 to 93.11 Hz.

3.5 T3 and T6

T3 and T6, traditionally, are mid-level and high-level tones with similar contours, but T3 should be higher than T6 in pitch level (Bauer & Benedict, 1997). The tonal pair T3 and T6 became slightly more distinct after the 1990s, while remained stable between the 1980s and the 1990s. Over time, T3 was consistently higher than T6, and both T3 and T6 displayed a falling tendency from the first and the fourth time points but retained level in the rest. Between the 1980s and the 1990s, T3 and T6 did not demonstrate significant changes both in pitch height and contour, persistently showing identical curves across time, with the pitch height difference about 2.5 Hz. Compared to the 1990s, the tonal pair after 2000 exhibited a slight change in the extent of decrease, along with a wider gap around 3.7 Hz in the final portion. Unlike the less than 3 Hz decline in the 1990s, T6 after 2000 dropped by more than 5 Hz, from 98.99 to 93.11. Meanwhile, T3 changed less obvious than T6. It decreased by 3.75 Hz in the 1990s whereas 4.91 Hz after 2000.
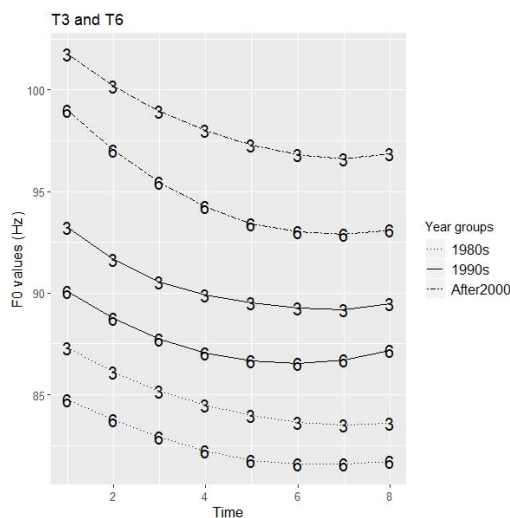

Figure 7. T3 and T6

3.6 T1 and its variant

T1 persistently showed falling trends across time. In LCY's speech, it continuously showed an evident decline tendency by 5.74 Hz in 1980s, 5.70 Hz in 1990s and 6.67 Hz after 2000. It also illustrated that the falling range increased across time. The phonetic realization of T1 thus seemed more similar to its high-falling variant, rather than the high-level one typical in HKC. It typically in HKC is a level tone placed at the top of one's normal pitch range, showing little changing pattern in pitch contour (Bauer & Benedict, 1997). Both the high-level and high-falling variants were used in the past (Chan, 1974), and speakers used them in free variation (Bauer, 1998; Kei, Smith, So, Lau, & Capell, 2002); whereas in more recent studies, only the high-level variant was described in HKC tonal system (Mok et al., 2013), suggesting the high-falling as more old-fashioned.
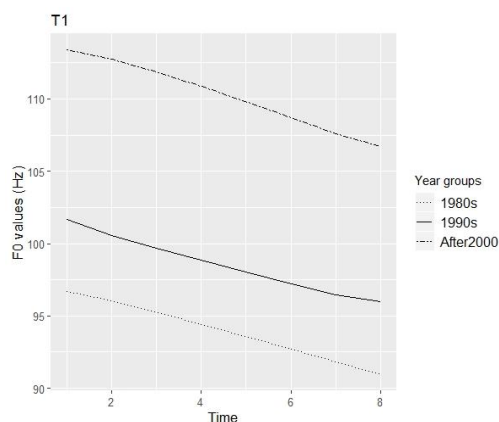

Figure 8. T1

4   Discussion

In general, the speaker changed his manner of tonal production over thirty years. The evidence again challenges the validity of apparent time hypothesis, suggesting that the speaker changed his accent in adulthood. In this regard, the findings are consistent with previous panel studies.

LCY's speech further exhibited sound changes in the opposite directions of the changes in progress in HKC: the six tone categories were preserved over time and increased in disparity, being distinct gradually. It is documented that though young speakers still retained six tone categories in production, some of them have been merging T2 with T5, T3 with T6 and T4 with T6, exhibiting much reduced pitch range (Mok et al.,

2013). Contrary to the communal trends, of the three tone pairs in LCY's speech, the two tones were separated gradually. The most apparent separation happened to T4 and T6, and T2 and T5 also improved in distinctiveness in pitch contour. T3 and T6 demonstrated increased differences in pitch height over time.

Therefore, Leung's journey of sound change seemed to go against community trends, showing greater conformity to conservative pronunciations over time. Primarily, the claim is anchored in the above observations of the three tone pairs, showing gradual separation rather than the merging progress as documented. Some side support for this claim can be obtained from T1, which, typically in HKC, is a high-level tone with little changing pattern. However, in LCY's speech, T1 was realized as its high-falling variant persistently, becoming more and more falling over time. As mentioned in Section 3.6, the high-falling variant of T1 is regarded as more old-fashioned in HKC. Hence, the evidence from T1 coincides with the gradual separation of other tone pairs, pointing to that LCY became more and more conservative in tone production.

A plausible explanation for the conservative pronunciations in LCY's speech is to relate his political career progression in the government to the authorities' efforts to restore standard forms, and to the presence of social stigma associated with nonstandard HKC pronunciations (i.e. "lazy speech"). The authorities have long been making efforts to correct pronunciation "mistakes": the language policies by the Education Department, the emergence of the committee for prescribing HKC pronunciations scholarly (廣州話審音委員會), and the media campaigns for promoting standard forms (as discussed in Chen, 2018). In their discourse, the adjective lazy (懶) could not only characterize the ways of pronunciation, but, according to Chen (2018), impose the language ideology the standard forms are of moral imperative, cultural superiority and political necessity. The standard forms thus greatly took on social meanings: an indicator of decency, education and power. Given the use of standard language variants is more often connected with competence, intelligence and status-related traits (Giles & Powesland, 1975), it is reasonable to speculate that with LCY's growth of political

engagement,[1] he adopted the standard forms in his public speech as a tool of image building among the citizens; whereas it is also plausible to say that LCY merely played a role in practicing the ideologies as a member of the authorities. Though his intention of language use is difficult to testify, the potential connections between the changes and the social meanings are unneglectable.

An alternative explanation is the variations in delivery style across the recordings of LCY's speech. For instance, the speaker could have improved his speech intelligibility consciously. Under this speculation, the changes in the individual's speech would be irrelevant to the sound changes but more related to hyper-articulation effects (Lindblom, 1990). The study indexed the degree of hyper-articulation by the speech rate, given previous studies reported that hyper-articulation led to wider pitch range, higher pitch level and slower speech rate (Freeman, 2010). The study thus hypothesized that if the increase in pitch range and level was caused by hyper-articulation, the speech rate should have been slow down simultaneously. However, in Section 3.2, the data illustrated that the speech rate of LCY increased from 1980s to 1990s and remained stable after 1990s, showing no correlation with pitch range and level and no evidence for hyper-articulation. Though there still exists the possibility for hyper-articulation, it is difficult to reconcile the theory with the results.

The potential aging effects on vocal tract should also be irrelevant to the observations in the study. A cross-sectional study measured the formant frequencies of spoken vowel by the young and the elder and reported the age-dependent decrease (Xue & Hao, 2003). The longitudinal results in male speakers are less consistent: F0 was reported to be unchanged (Verdonck-de Leeuw & Mahieu, 2004), to decline (Reubold, Harrington, & Kleber, 2010) and to incline occasionally (Harnsberger, Shrivastav, Brown, Rothman, & Hollien, 2008). Though a more sizable body of literature supports the age-dependent decrease in F0, suppose the pitch level increase by LCY was influenced by changes in physiological configuration, it still

cannot provide reasons for the expanded pitch ranges and increased distinctiveness of tones.

5 Conclusion

Taken together, Leung's journey of sound change seemed to change against community trends, showing greater conformity to conservative pronunciations over time, especially since the 1990s. The results showed the increased pitch level and the expanded pitch range over time, with no correlations with speech rates; the increased distinctiveness in the three tonal pairs over time; the high-falling variant of T1 and the growth of its falling extents over time. We discussed the results in the framework of changes across the lifespan, relating to LCY's career progression, the government's efforts to standardize HKC and the presence of social stigma about nonstandard HKC pronunciations. The aggregated evidence also suggested the irrelevance of stylistic variations and aging effects. Overall, the study presents another possibility beneath the communal trends: not bound to follow the trends, an individual could change in an opposite direction across his lifespan. The discrete cosine transform could be further implemented for analyzing tonal curves in the study (Ahmed, Natarajan & Rao, 1974). More real time data need to be collected and combined with apparent time evidence, so that a fuller picture of sound change could be depicted.

References

Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete cosine transform. *IEEE Transactions on Computers*, *100*(1), 90–93.

Bailey, G. (2002). Real and apparent time. *The Handbook of Language Variation and Change*, 312–332.

Bauer, R. S. (1998). Hong Kong Cantonese Tone Contours. In S. Matthews (Ed.), *Studies in Cantonese Linguistics* (pp. 1–34). Hong Kong: Linguistic Society of Hong Kong.

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Berlin: Mouton de Gruyter.

Bigi, B., & Hirst, D. (2012). SPeech Phonetization Alignment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody. In *Speech Prosody* (pp. 1–4).

---

[1] In 1970s, he was a surveyor and businessman; starting from 1980s, he gradually became involved in political affairs and became a member of Hong Kong´s Basic Law Consultative Committee in 1985; in 2012 he was officially appointed as the Chief Executive of Hong Kong.

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Retrieved April 3, 2018, from http://www.praat.org/

Chan, Y.-Y. F. (1974). *A perceptual study of tones in Cantonese*. Centre of Asian Studies, University of Hong Kong.

Chen, K. H. Y. (2018). Ideologies of Language Standardization: The case of Cantonese in Hong Kong. In *The Oxford Handbook of Language Policy and Planning*. UK: Oxford University Press.

Freeman, V. (2010). Using hyperarticulation to quantify interaction between discourse functions. *The Journal of the Acoustical Society of America*, *128*(4), 2457.

Fung, R., & Bigi, B. (2015). Automatic word segmentation for spoken Cantonese. In *Oriental COCOSDA held jointly with 2015 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE), 2015 International Conference* (pp. 196–201). IEEE.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation.* Academic Press.

Harnsberger, J. D., Shrivastav, R., Brown Jr., W. S., Rothman, H., & Hollien, H. (2008). Speaking Rate and Fundamental Frequency as Speech Cues to Perceived Age. *Journal of Voice*, *22*(1), 58–69.

Harrington, J., Palethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association*, *30*(1–2), 63–78.

Hubert, M., & Van der Veeken, S. (2008). Outlier detection for skewed data. *Journal of Chemometrics*, *22*(3-4), 235–246.

Kei, J., Smith, V., So, L. K. H., Lau, C. C., & Capell, K. (2002). Assessing the accuracy of production of Cantonese lexical tones: a comparison between perceptual judgement and an instrumental measure. *Asia Pacific Journal of Speech, Language and Hearing*, *7*(1), 25.

Kwan, T. (2003). Chinese Character Database: With Word-formations Phonologically Disambiguated According to the Cantonese Dialect. Hong Kong: Research Centre for Humanities Computing, CUHK. Retrieved from http://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/

Kwon, S. (2018). Phonetic and phonological changes of Noam Chomsky: A case study of dialect shift. *American Speech*, *93*(2), 270–297.

Labov, W. (1999). *Principles of linguistic change, volume 1: Internal factors* (Vol. 36). Oxford: Blackwell.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403–439). Dordrecht: Springer Netherlands.

Mok, P. P. K., Zuo, D., & Wong, P. W. Y. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change*, *25*(3), 341–370.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, *52*(7), 638–651.

Sankoff, G., & Blondeau, H. (2007). Language change across the lifespan: /r/ in Montreal French. *Language*, *83*(3), 560–588.

Shapp, A., LaFave, N., & Singler, J. V. (2014). Ginsburg v. Ginsburg: A longitudinal study of regional features in a Supreme Court justice's speech. *University of Pennsylvania Working Papers in Linguistics*, *20*(2), 149–158.

To, C. K. S., Mcleod, S., & Cheung, P. S. P. (2015). Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech sound assessment. *Clinical Linguistics & Phonetics*, *29*(5), 333–353.

Traunmüller, H., & Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. *Unpublished Manuscript*.

Verdonck-de Leeuw, I. M., & Mahieu, H. F. (2004). Vocal aging and the impact on daily life: a longitudinal study. *Journal of Voice*, *18*(2), 193–202.

Xu, Y. (2013). ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. In *Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)* (pp. 7–10). Aix-en-Provence, France.

Xue, S. A., & Hao, G. J. (2003). Changes in the Human Vocal Tract Due to Aging and the Acoustic Correlates of Speech ProductionA Pilot Study. *Journal of Speech, Language, and Hearing Research*, *46*(3), 689–701.