Citation information:
Tse, Holman (2016). Variation and change in Toronto heritage Cantonese. *Asia-Pacific Language Variation*, *2*(2), 124–156. https://doi.org/10.1075/aplv.2.2.02tse

Published online: March 24, 2017

Variation and change in Toronto heritage Cantonese: An analysis of two monophthongs across two generations

Holman Tse

University of Pittsburgh

This paper presents the first sociophonetic study of Cantonese vowels using sociolinguistic interview data from the Heritage Language Variation and Change in Toronto Corpus. It focuses on four allophones [iː], [ɪk/ɪŋ], [uː], and [ʊk/ʊŋ] of two contrastive vowels /iː/ and /uː/ across two generations of speakers. The F1 and F2 of 30 vowel tokens were analyzed for these four allophones from each of 20 speakers (N = 600 vowel tokens). Results show inter-generational maintenance of allophonic conditioning for /iː/ and /uː/ as well as an interaction between generation and sex such that second-generation female speakers have the most retracted variants of [ɪk/ɪŋ] and the most fronted variants of [iː]. This paper will discuss three possible explanations based on internal motivation, phonetic assimilation, and phonological influence. This will illustrate the importance of multiple comparisons (including inter-generational, cross-linguistic, and cross-community) in the relatively new field of heritage language phonology research.

Keywords: heritage languages, bilingualism, sociophonetics, contact linguistics, Cantonese/ Chinese language--Canada

1. Introduction

This paper presents results from the first sociophonetic study of vowel variation and change in Toronto Heritage Cantonese (HCAN) by focusing on the production of two vowels and their four allophonic variants. It contributes to variationist research in two ways. First, it examines the vowel system of a heritage language (HL). Second, it focuses on Cantonese, a variety of Chinese that has been under-examined in terms of vowel variation. The research questions are as follows: (1) Is allophonic conditioning maintained across two generations of Cantonese speakers in Toronto, Canada for the vowels /iː/ and /uː/? (2) Is there evidence for phonetic variation in the allophones of /iː/ and /uː/ based on generational group or sex? (3) What best accounts for observed variation: internal motivation, phonetic assimilation, or phonological influence?

I follow Nagy's (2011) definition of a HL speaker in Canada as "anyone who is a mother tongue speaker of a language identified with their heritage other than French or British" (p. 68). The social setting in which a HL develops is one that involves a sociolinguistic transition from an immigrant generation (GEN 1) to a generation born in a different country in which a different language is a socially and politically dominant language (GEN 2). Under such a social context, we might expect contact-induced change. The small, but growing literature on HL vowels, however, shows a lack of change in terms of phonological contrasts. Yet, these studies also show low-level phonetic changes influenced by both phonetic similarity with vowels in the dominant language and by phonological considerations (Chang, Yao, Haynes, & Rhodes, 2011; Godson, 2004; Ronquest, 2013). One question that has not been previously addressed is whether the dominant language can have any effect on the acoustic production of allophones in a HL.

This paper addresses this question using data from the HerLD (Heritage Language Documentation) Corpus, which was developed as part of the Heritage Language Variation and Change (HLVC) in Toronto Project (Nagy, 2011). This corpus includes sociolinguistic interviews from speakers of eight[1] HLs spoken in the Greater Toronto Area (GTA) all collected using the same methodology. This standardized methodology facilitates analyses based on three major sets of comparisons (inter-generational, cross-linguistic, and cross-community), which are all important for addressing theoretical questions about contact-induced change. The present study involves an inter-generational comparison. Results show maintenance of allophonic conditioning of /iː/ and /uː/ based on velar context across two generations. Though GEN 2 female speakers produce significantly more fronted variants of [iː] than do other speaker groups, results also show inter-speaker variation among both generational groups that could be due to multiple factors including pre-existing variation in Hong Kong Cantonese and different ways in which the Toronto English (TOENG) vowel system could influence HCAN vowels. This uncertainty illustrates the importance of future research incorporating both a cross-linguistic comparison using normalized acoustic data from TOENG and a cross-community comparison using normalized acoustic data from Hong Kong Cantonese.

2. Background

2.1 Heritage language phonology research

In contrast to much of the literature on HLs, I adopt a variationist perspective by avoiding assumptions about linguistic deficit as illustrated by the title of Montrul's book *Incomplete Acquisition in Bilingualism* (2008) and Polinsky's (2011) definition of HLs as languages "spoken by early bilinguals, simultaneous or sequential, whose home language (L1) is severely restricted

because of insufficient input" (para. 1). Much of the literature has also focused on low-proficiency speakers. This paper examines the speech of only those fluent enough to carry on a one-hour long conversation in Cantonese as required for participation in the HLVC Project (Nagy, 2011). Adopting a variationist approach also means recognizing the possibility that socially-conditioned variation may exist among both GEN 1 and GEN 2 speakers. Thus, any inter-generational differences observed may not necessarily be due to influence from the dominant language.

Although HL research in general is a relatively new field, much of the existing work has focused on morpho-syntax rather than on phonetics or phonology (Polinsky & Kagan, 2007). Only a few published studies have focused specifically on vowels. These studies discuss different sources of phonetic features of HL speech. One source is phonetic similarity with the dominant language. Godson (2004), for example, shows that HL western Armenian speakers in southern California are influenced by English in their production of western Armenian /i/, /ɛ/, and /a/ but not in their production of /u/ and /o/. This is because /i/, /ɛ/, and /a/ have phonetically similar counterparts in California English while western Armenian /u/ and /o/ lack similar counterparts.

Another possible source is the transfer of phonological rules from the dominant language. Ronquest (2013) discusses this possibility for HL Spanish speakers in Chicago who produce more centralized vowels in unstressed than in stressed syllables. This could be influenced by vowel reduction in English. Another possibility discussed by Ronquest (2013) is the influence of non-standard features of GEN 1 speech. The lack of acoustic data on GEN 1 speakers in the study could possibly explain why the direction of centralization observed does not match the production of English unstressed schwa.

Perhaps the most intriguing possibility is the role that cross-linguistic phonological considerations can play. Chang, Yao, Haynes, and Rhodes (2011) compare different groups of Mandarin-English bilinguals and show that HL speakers produce Mandarin /u/ and Mandarin /ou/ further back (lower F2) than the L2 Mandarin group. While the L2 Mandarin speakers may be influenced more by the relatively high F2 of English /u/ and /ou/ (phonetic assimilation), HL Mandarin speakers may be influenced more by a need to maintain a phonological contrast between /u/ and /y/. Overall, the HL speakers examined produced the greatest average F2 difference between Mandarin /u/ and English /u/. For some HL speakers, this even meant lower F2 than native monolingual Mandarin speakers. Early exposure to two phonological systems may account for why HL speakers produce greater cross-linguistic and language-internal distinctions (in both languages) than L2 speakers.

Though these three studies discuss different possible sources of HL phonetic features, they all show contrast maintenance. While phonetic influence from the dominant language can sometimes play a role as in the western Armenian case, the need to maintain a phonological contrast, as in the Mandarin /u/ versus /y/ case, can sometimes override the influence of phonetic similarity. The fact that phonological contrasts are maintained corroborates the widespread impression among researchers and language teachers that HL phonology is 'native-like' in contrast to HL morpho-syntax (Polinsky & Kagan, 2007). Contrast, however, is only one aspect of phonology. Allophonic conditioning is another and is one that is not considered in previous HL research. It may have implications for sociophonetic differences between GEN 1 and GEN 2 speakers.

2.2 Socio-historical background

The community under investigation in this paper is Toronto's Cantonese speaking community, which is one of the largest outside Asia. With over 177,000 speakers, Cantonese is almost tied with Italian (178,000 speakers) as the second most widely spoken language after English (Statistics Canada, 2012, Figure 4.2). This is the result of several major waves of immigration. The first occurred after the easing of immigration laws in the 1960s (Thompson, 1989). Major waves also came in the 1980s and 1990s and were largely motivated by the political uncertainty of Hong Kong's future especially with the 1997 handover of Hong Kong from a British colony to the government of the People's Republic of China (Li, 2005). Though a Cantonese speaking community has been present in Toronto since the 19th Century, the community had been relatively small compared to other North American Chinese communities until after the 1960s. The earlier immigrants came largely from the Siyi (the four counties: Taishan, Kaiping, Enping, Xinhui) area of Guangdong province (Hsu, 2000; Thompson, 1989), where a continuum of dialects are spoken. These dialects have been described as mutually unintelligible with the prestige variety spoken in Hong Kong (Yue-Hashimoto, 1972).

Siemiatycki, Rees, Ng, and Rahi have also noted that the large size and affluence of Toronto's Chinese community has made it possible "to promote an impressive commercial, media, and marketing presence" (2003, p. 408). This includes having five Chinatowns, three Chinese- language daily newspapers, two Chinese-language television stations, Chinese editions of several English-language magazines, and a large number of businesses that advertise in Cantonese. Hence, the opportunities for use and exposure to Cantonese outside of the home context are quite the opposite of 'severely restricted' as suggested by Polinsky's (2011) definition of HL speakers. Previous research in the Toronto Chinese community has shown that

many GEN 2 Chinese-Canadians are bilingual in English and Cantonese (Hoffman & Walker, 2010). Man (2006) has also reported strong ethnolinguistic vitality within the Toronto Chinese community. Although cases of attrition on an individual level do exist, they appear to be less common than observed in other HL research contexts especially in the US.

Toronto, Canada as a whole has also been widely described as a city with a social environment that supports sustained multilingualism. Part of this can be attributed to the Canadian government's adoption of a multiculturalism policy in 1971 (Siemiatycki et al., 2003). In the past few decades, this policy has had an especially profound impact on life in Toronto and on how residents describe their city. A *Toronto Star* article, for example, describes the Greater Toronto Area (GTA) as "unofficially but inexorably … approaching omni-lingualism" and as a place where one can hear residents speaking many different languages *by choice* suggesting that societal pressures to shift to exclusive use of English, the lingua franca, are minimal (Taylor, 2007, p. ID06). This article also describes how different parts of the GTA are "a conurbation of neighbourhoods [sic], rather than ghettos" suggesting a lack of social stigma attached to ethnic minority identity. Toronto is both a city with a large concentration of speakers of non-official languages and a city with social conditions supportive of HL maintenance making it an ideal place to research HL variation and change.

2.3 The Cantonese vowel system

Cantonese has a typologically large vowel inventory that includes eight contrastive monophthongs (see **Error! Reference source not found.**) and 11 contrastive diphthongs. The vowels, /iː/ and /uː/, which are the focus of this paper, are pronounced as [ɪ] and [ʊ] respectively when followed by a velar consonant [k] or [ŋ] and as [iː] and [uː] respectively elsewhere (Yue-

Hashimoto, 1972). Some research has suggested that the actual pronunciation has lowered to [e] and [o] respectively at least for some speakers in Hong Kong.

Lee (1983), for example, shows inter-speaker variation among three speakers from Hong Kong and three speakers from Guangzhou, China. Two of the Hong Kong speakers and one of the Guangzhou speakers lower [ɪk/ɪŋ][2] to a F1/F2 range that overlaps with their pronunciation of the vowel /ɛ/ while the remaining three speakers have quite distinct pronunciations. All six speakers show considerable overlap between [ʊk/ʊŋ] and /ɔ/. More recently, Zee (2003) shows considerable overlap between [ɪk/ɪŋ] and /ɛ/ as well as overlap between [ʊk/ʊŋ] and /ɔ/ based on F1/F2 means among 50 male and 50 female Hong Kong speakers between the ages of 18 and 21. The twenty-year time difference between these two studies suggests that the lowering of [ɪk/ɪŋ] and [ʊk/ʊŋ] are, in fact, sound changes that have occurred or are in progress in Hong Kong Cantonese.

The question of variation has not been a concern in most of the existing literature on Cantonese vowels. Instead, much work on Cantonese vowels has focused on normative descriptions that ignore the possibility of socially patterned variation. Bauer and Benedict (1997), for example, review almost a dozen descriptions of Hong Kong Cantonese vowels. They discuss a controversy over whether these allophones should be transcribed [ɪ] and [ʊ] as assumed in this paper or as [e] and [o]. Bauer and Benedict (1997, p. 45) argue that the reason for this discrepancy is that these vowels are phonetically short monophthongs that combine acoustic properties of both [ɪ] and [e] and of both [ʊ] and [o]. This results in conflicting impressionistic descriptions. They propose that these allophones should instead be transcribed as [e^{j}] and [o^{w}] respectively. Similarly, Zee (2003) also attributes transcription discrepancies to the acoustic characteristics of these allophones. Since these allophones are produced in an environment in

which they are phonetically short, there is a mismatch between production and perception. This mismatch, he says, explains why different authors have transcribed these allophones in different ways. Again, the focus is on transcription of a hypothetical norm rather than on variation.

Given Lee's (1983) study, however, it is clear that inter-speaker variation could also account for transcription discrepancies. Lee (1983) is also one of the very few studies that recognizes the possibility of synchronic variation in this part of the vowel system. While variation in consonants (see Matthews & Yip, 2011, pp. 35–37 for a list of features and discussion) and tone (Bauer, Cheung, & Cheung, 2003; Mok, Zuo, & Wong, 2013) in Cantonese have previously been investigated, vowels remain a surprisingly under-researched part of the Cantonese sound system in sociolinguistics literature.

3. Research questions and hypotheses

3.1 Research questions

The specific research questions stated in the introduction are repeated as follows:

1) Is allophonic conditioning of /iː/ and /uː/ maintained across two generations of Toronto HCAN speakers?

2) Is there evidence for low-level phonetic variation in the allophones of /iː/ and /uː/ based on generational group or sex?

3) What best accounts for observed variation: internal motivation, phonetic assimilation, or phonological influence?

The first two questions will be addressed by determining whether or not velar context, generational group, and sex are significant predictors of variation and by determining whether or not there is an interaction between these variables. The third question will require interpreting

what this interaction means. Section 3.2 presents three sets of hypotheses based on findings from previous research on HL phonology (Section 2.1) and from previous research on Cantonese vowels (Section 2.3). These hypotheses are summarized in Table 2.

3.2 Hypotheses

3.2.1 *Internal motivation*

The first set of hypotheses is based on pre-existing variation found in Hong Kong Cantonese. Section 2.3 discussed the lowering of the pre-velar allophones [ɪk/ɪŋ] and [ʊk/ʊŋ] as recent sound changes in Hong Kong. If Lee (1983) found variation in the lowering of these allophones in the 1980s, then it would not be a surprise to also find variation among GEN 1 speakers since all GEN 1 speakers were born before the 1980s. GEN 2 speakers could also be variable and could even be pushing the change towards completion. Lau (2003) has argued that the lowering of these allophones is part of a longer term chain shift that supports one of Labov's (1994) principles for vowel chain shifts, namely Principle II— short vowels fall. Thus, if both [ɪk/ɪŋ] and [ʊk/ʊŋ] are lowered in Toronto HCAN, the lowering (or increase along the F1 dimension) could be due to internal motivation[3]. The movement from [ɪ] to [ɛ] may also involve some retraction along the F2 dimension.

3.2.2 *TOENG influence through phonetic assimilation*

The second hypothesis is that TOENG vowels affect HCAN vowels through phonetic assimilation as has been observed for /i/ and /ɛ/ in HL western Armenian (Godson, 2004). Under this hypothesis, the HCAN vowels would become increasingly similar to the most similar

TOENG vowels. In this case, the four Cantonese allophones in question share phonetic similarity with four contrastive TOENG vowels.

For HCAN [uː] and [ʊk/ʊŋ], the closest TOENG counterparts are /u/ and /ʊ/ respectively, which are both fronted as in many dialects of English. If there is influence from TOENG due to phonetic assimilation, then we would expect [uː] and [ʊk/ʊŋ] to also front (higher F2). For HCAN [iː], the closest TOENG counterpart is /i/. Since these are point vowels in each language, there is no a priori reason to believe that they are significantly different from each other. If this proves to be the case, the phonetic assimilation hypothesis would predict no change for [iː].

The fourth allophone, [ɪk/ɪŋ], is most phonetically similar to TOENG /ɪ/, a vowel that has also been subject to a major sound change. This sound change is part of the Canadian Vowel Shift, which involves the lowering and retraction of three front lax vowels including /ɪ/ (Clarke, Elms, & Youssef, 1995). Though recent work on TOENG has shown stability in /ɪ/ and has also shown retraction rather than lowering as the primary direction of change (Roeder & Jarmasz, 2010), it seems possible that GEN 2 HCAN speakers may be influenced by the already lowered and retracted /ɪ/. If this turns out to be the case, we would expect GEN 2 HCAN speakers to have lowered and retracted [ɪk/ɪŋ] compared to GEN 1 speakers. This turns out to be the same direction of movement predicted by the internal motivation hypothesis. However, results will show patterns that may be better accounted for based on a third hypothesis.

3.2.3 *TOENG phonological influence*

The third hypothesis to account for variation is also one attributable to influence from TOENG but through phonological influence rather than through phonetic assimilation. This hypothesis builds on Chang et al.'s (2011) study showing that HL bilingual speakers have a greater acoustic

difference between phonetically similar vowels in their two languages compared to L2 bilingual speakers. This results from early exposure to two phonological systems and the pressure to keep the two systems distinct. Chang et al. state that "this pressure may even push the categories further apart than they need to be" (2011, p. 3975) so that the outcome is that some HL speakers in their study produced Mandarin /u/ further back than monolingual Mandarin speakers. Thus, one strategy of keeping two vowel systems apart could be in making point vowels in one language more peripheral. In the study presented in this paper, two of the HCAN allophones are also point vowels, namely [iː] and [uː]. One prediction about how these two vowels might change under influence from TOENG is that both of these vowels become more peripheral. This could mean either fronting (higher F2) or raising (lower F1) of [iː] and either backing (lower F2) or raising (lower F1) of [uː].

<Insert Table 2 HERE>

4. Data

The data examined in this paper include recorded interviews (in .wav format) of 20 Cantonese speakers from the HerLD corpus[4] (Nagy, 2011). This sample includes an equal number of male and female speakers as well as equal representation from each of the two generation groups. As required by the HLVC Project guidelines, all of the GEN 1 speakers in the corpus grew up in Hong Kong, moved to the Greater Toronto Area (GTA) as adults, and have lived in the GTA for at least 20 years. Most also speak at least some English though age of acquisition and proficiency levels vary. This includes variable exposure to a British variety of English through formal schooling in Hong Kong[5] and additional exposure to a different variety of English in Toronto. What is most important for the purpose of this paper is that all GEN 1 speakers speak Cantonese

as their mother tongue and that none reported acquiring TOENG as children. GEN 1 speakers form the baseline group for comparison in this study.

For GEN 2 speakers, on the other hand, both English and Cantonese were acquired at an early age in a societal setting in which TOENG is the dominant language. GEN 2 speakers include those who have parents that meet the qualifications for GEN 1 speakers. In addition, GEN 2 speakers must have grown up in the GTA. This means that they were either born in the GTA or moved to the GTA before the age of three[6] and have since lived continuously in the GTA. Consequently, all GEN 2 speakers speak English as a dominant language[7].

The 20 speakers examined in this paper are listed in Table 3 according to their speaker code. In addition to .wav file recordings, the corpus also includes time-aligned transcriptions of the data completed by native speakers using the program ELAN (Sloetjes & Wittenburg, 2008). The Cantonese data was transcribed using the Jyutping Romanization System. To deal with code switching and code mixing with English, the general rule of thumb was to transcribe in English orthography if the pronunciation was more English-like and in Jyutping if the pronunciation was more Cantonese-like. Only words that are unambiguously Cantonese are analyzed in this paper.

<Insert Table 3 HERE>

5. Method

5.1 Overview

The first step was to export a textgrid from the ELAN transcript readable by the phonetics analysis software, Praat (Boersma & Weenink, 2014), which was also used to collect F1 and F2 measurements of a sample of vowels. A total of 75 vowel tokens were recorded on a spreadsheet for each speaker. With a total of 20 speakers, this amounted to a grand total of 1500 vowel

tokens. These 75 vowel tokens for each speaker included tokens from five vowel categories (/iː/, /uː/, /ɛː/, /ɔː/, and /aː/), and two contexts (pre-velar and non-pre-velar).

<Insert Table 4 HERE>

Table 4 shows the composition of the 75 tokens collected for each speaker. Fifteen tokens were collected for each vowel category. This included 10 tokens in open syllable context and five in pre-velar context[8]. An exception was made for /uː/ due to the low type and token frequency of /uː/ in open syllable contexts. /uː/ occurred far more frequently in pre-velar than in other contexts. For /uː/, this distribution was 10 tokens in pre-velar context and five tokens in non-pre-velar contexts[9]. An exception was also made for /aː/, which includes only open syllable context due to its low overall token frequency in pre-velar contexts. Since /aː/ is not the focus of this study, this was not a concern. Yet, /aː/ was still included as a point vowel for the purpose of normalization. The other vowels /ɛ/ and /ɔ/ will be used as reference points for describing the trajectory of change of /iː/ and /uː/ in pre-velar context. They will be analyzed in more detail in a separate study.

Finally, to control for the potential effects of tone, only words produced with the high-falling tone (also known as Tone 1) were included with an exception made for /uː/ due to low type and token frequency. The words chosen were the first five that occurred in the recording after the 15-minute point. Tokens were also taken from the first 15 minutes if there were not enough tokens elsewhere in the recording. If these exceptions were not made for /u/, some speakers would have had zero tokens of [uː]. Ultimately, tone was found to have no significant effect on formant frequencies.

5.2 Procedure

The formant measurement procedures began with the exported Praat textgrid. The exported transcript tier was reviewed word-by-word for tokens of interest. Once a relevant token was identified, the boundaries of the word and vowel were added. Measurements were taken by selecting the entire vowel and using Praat's "get formant" command to calculate both the F1 and F2 means. Co-articulation effects were minimized by selecting only the steady state portion of vowels. The default settings were set to 5500 Hz for the maximum formant, five formants, and a window length of 0.025 seconds[10]. These settings were adjusted in cases of obvious errors.

5.3 Normalization and subsequent steps

After recording measurements for these 1500 tokens, the next step was to normalize the vowel measurements using the NORM vowel normalization suite (Thomas & Kendall, 2007). The Lobanov technique was selected and the output was scaled to Hertz values. Finally, the last step was to upload the normalized data to Rbrul (Johnson, 2009) for statistical analysis of vowel formant measurements. **Table** 5 shows the independent variables examined. The dependent variable was either F1 (as an inverse acoustic correlate of vowel height) or F2 (as an acoustic correlate of vowel frontness). Separate analyses were conducted for each formant frequency of each vowel. The independent variables included both random effects (speaker and word) and fixed effects, which included external (generation and sex) as well as linguistic factors (following velar consonant and tone)[11].

To determine the extent to which male and female speakers from different generation groups may differ in their production of each allophone, a factor group was also created combining three categorical variables (generation, sex, and following velar consonant) for a total

of 8 possible values. Whenever this factor group was included in a run, generation, sex and velar consonant were excluded as distinct variables in the modeling.

<Insert **Table** 5 HERE>

6. Results

6.1 Overall results

The overall results of the formant measurements are represented in the charts shown in Figure 1 and Figure **2**2. Both of these charts show plots of the Lobanov normalized mean F1 and F2 values for each of the four factor groups (GEN 1 female, GEN 1 male, GEN 2 female, GEN 2 male) for each of the four allophones ([iː], [ɪk/ɪŋ], [uː], [ʊk/ʊŋ]) under examination. Figure 1 shows variability in F1/F2 means across speaker groups. For example, GEN 2 female speakers show the greatest difference in means between these two allophones along the F2 axis while GEN 1 female speakers show the smallest difference. Yet, the F1/F2 means of [iː] and [ɪk/ɪŋ] are still clearly separated from each other with [ɪk/ɪŋ] consistently lower (higher F1) than [iː] across all speaker groups.

< Insert Figure 1 HERE>

Figure **2**2 shows a clear distinction between the two allophones of /uː/ as well as inter-group variability. The F1 means for all speaker groups for [ʊk/ʊŋ] are lower in the vowel space (higher in F2) than any of the means for [uː]. Unlike in Figure 1, there is no clear pattern in terms of whether or not any speaker group shows greater phonetic distinction in these allophones than other groups.

[Figure **2**2 HERE]

Figures 3-6 are vowel plots for each of the four speaker groups. Each plot includes the mean ± 1 standard deviation for each allophone. This is represented in the form of an ellipsis to make it

easy to visualize the relative degree of overlap between different allophones. Each plot also uses the same Hertz range on the x and y axes to facilitate comparisons across speaker groups. **Error! Reference source not found.**3 shows that [ɪk/ɪŋ] is quite distinct from /ɛ/ for GEN 1 female speakers. In contrast, GEN 1 male speakers show overlap between [ɪk/ɪŋ] and /ɛ/ (see Figure 4). GEN 2 speakers as a whole also show overlap between [ɪk/ɪŋ] and /ɛ/ although the overlap is greater among GEN 2 female speakers (see Figure 5) than it is for their male counterparts (see Figure6).

Another major difference across speaker groups is in the allophones of /iː/. GEN 2 female speakers show no overlap between [iː] versus [ɪk/ɪŋ] in terms of mean ± 1 standard deviation (Figure 5). While GEN 2 male speakers do show some overlap (Figure6), the overlap appears to be relatively small in contrast to that of GEN 1 speakers. **Error! Reference source not found.**3 and Figure 4 show that the mean ± 1 standard deviation of [ɪk/ɪŋ] extends much closer to the mean F1/F2 of [iː] for both GEN 1 male and female speakers than it does for either of the GEN 2 speaker groups.

We can also find variation in patterns of overlap for the back vowel allophones. [uː] appears to be distinct from [ʊk/ʊŋ] for all speaker groups though to different degrees. These two allophones appear to be the most distinct for GEN 2 female speakers (Figure 5) and the least distinct for GEN 1 female speakers (**Error! Reference source not found.**3). [ʊk/ʊŋ] also overlap to varying degrees with allophones of /ɔ/. While GEN 1 male (Figure 4) speakers have the least amount of overlap, it does not seem clear which speaker group has the most overlap.

<Insert **Error! Reference source not found.**3, Figure 4, Figure 5, Figure6 HERE>

6.2 Results for individual vowels

6.2.1 *The vowel /iː/*

Table 6 shows the best step-down model of F1 as the dependent variable for /iː/. The only significant factor is velar context ($p < .001$). The mean F1 of /iː/ before velar consonants is 427 Hz and 373 Hz in open syllable context. A higher F1 before velar consonants means that /iː/ is lower in the vowel space in this environment. In this case, the step-up and step-down models match. This supports the observations made in Section 6.1 about the acoustic distinctiveness of [iː] versus [ɪk/ɪŋ]. For the F2 of /iː/, on the other hand, none of the independent variables showed statistical significance. This does not rule out the possibility of significant predictors in the two allophones of this vowel.

<Insert **Table** 6 HERE>

To address whether or not there are any significant differences in how male and female speakers from the two generational groups produce the two allophones of /iː/, **Table** and **Table** 8 show the results of mixed effects models that put generation, sex, and velar context together as a single factor group. **Table** 7 shows the best step-down model for the F1 of /iː/ while **Table** 8 shows the best step-down model for the F2. The GEN×Sex×Velar factor group is significant for both F1 ($p = 0.00312$) and F2 ($p < 0.000101$).

To address which speaker groups are significantly different from which other groups, a set of Tukey Post-Hoc tests were run for both F1 and F2. The results for F1 show significance for pairings involving [iː] and [ɪk/ɪŋ] for each speaker group (**Table** 9). This shows a very clear separation between [iː] and [ɪk/ɪŋ] across all speaker groups. The coefficient rankings in **Table** 7 show further evidence of this separation. All of the positive coefficient values as well as the

highest means occur with [ɪk/ɪŋ] while all of the negative values (as well as the lowest means) occur with [iː]. Higher F1, thus, co-occurs with [ɪk/ɪŋ] while lower F1 co-occurs with [iː].

<Insert **Table** 7, **Table** 8 HERE>

In contrast, the Post-Hoc results for F2 pairings involving [iː] and [ɪk/ɪŋ] did not show across-the-board significance. We can visualize why this may be the case by returning to Figure 1. Overall, there is less of a separation in allophones along the F2 axis. The notable exception is for the GEN 2 female group. The GEN 2 male group also shows some separation along the F2 axis, though not as much as the GEN 2 female group. **Table** 8 shows the GEN 2 female group on opposite ends in terms of coefficient values and means. The GEN 2 female group has the highest F2 for [iː] and the lowest F2 for [ɪk/ɪŋ]. The results from the Tukey Post-Hoc test shown in **Table** 9 support the observation that GEN 2 females have significantly higher F2 with [iː] and lower F2 with [ɪk/ɪŋ] ($p < 0.001$).

<Insert **Table** 9 HERE>

Table 10 shows the results of Tukey Post-Hoc tests addressing whether or not there are significant inter-group differences in the production of [iː] while **Table 2**1 shows the results for [ɪk/ɪŋ]. Both tables include a list of every possible pairing between the four different speaker groups. The Post-Hoc tests show no statistically significant differences in the F1 values of [iː] based on speaker group. For F2, on the other hand, there is statistical significance only for pairings involving the GEN 2 female group. In fact, every possible pairing between GEN 2 female speakers and other groups shows statistical significance. With the highest F2 values, this means that GEN 2 female speakers show the greatest tendency of fronting [iː].

Table 21 also shows the lack of significant inter-group differences in the F1 of [ɪk/ɪŋ]. The only statistically significant result shown in this table is in the F2 between GEN 1 and GEN

2 female speakers ($p < 0.042$). This suggests that there is a continuum in terms of F2 values with GEN 1 female speakers on one end, GEN 2 female speakers on the other end, and other speakers somewhere in between with overlapping distributions. On the one hand, this shows that the GEN 2 female group is a group that stands out in contrast to all other groups. Yet, on the other hand, this could be a false positive. If we look closely at the r^2 values in **Table** 8, we can see evidence for considerable inter-speaker variability along the F2 axis. **Table** 8 shows a r^2 value of 0.21 for the random effects. This is almost twice as large as the r^2 value for the GEN×Sex×Velar factor group ($r^2 = 0.126$). This indicates that the different factor groups account for a smaller percentage of the variation than do factors such as Speaker and Word. This shows that there is both inter-speaker and word-based variation in F2 that cannot be accounted for based only on Generation group or Sex.

< Insert Table 10, Table 21 HERE >

6.2.2 *The vowel /uː/*

For the vowel /uː/, the first set of results (**Table 3**2) comes from the best step-down model of F1 with generation, sex, and velar contexts as separate independent variables. As is the case for the F1 of /iː/, the only significant fixed effect is for velar context ($p < .001$). Velar consonants condition higher F1 for /uː/ just as they do for the F1 of /iː/. Tone is also included in this model since it was not controlled as a variable in selecting tokens of /uː/ (see Section 5.1). Ultimately, tone did not show a main effect in any of the models in which it was included. Also as is the case for /iː/, none of the independent variables tested showed a main effect for F2.

<Insert **Table 3**2 HERE>

Table 43 and **Table 5**4 present the best step-down models for F1 and F2 respectively with generation, sex, and velar context grouped together as a single factor group. These results

address the question of inter-speaker variation in the pronunciation of allophones. This factor group came out significant in both models. Also, in both cases, the best step-up and step-down models match. We can also see a clear separation in the coefficient rankings based on velar context. [ʊk/ʊŋ] for each speaker group has higher F1 while [uː] has lower F1. This shows that the pre-velar allophones are lower (higher F1) in the vowel space for all speaker groups. Post-Hoc tests confirm a clear separation between [uː] and [ʊk/ʊŋ] along the F1 axis, but not along the F2 axis. None of the comparisons between different groups for [uː] nor for [ʊk/ʊŋ] were significant. There is, hence, lack of evidence for significant inter-group differences in the production of these two allophones. **Table 6**5 summarizes all the Post-Hoc results discussed in this section.

<Insert **Table 4**3, **Table 5**4, **Table 6**5 HERE>

7. Discussion

7.1 Maintenance of allophonic conditioning

The first research question of this paper is about whether or not Toronto HCAN speakers exhibit allophonic variation in /iː/ and /uː/ based on pre-velar context. All of the results presented in Section 6 support an affirmative response to this question. Velar context was the only significant fixed effect in the model presented in **Table** 6. The F1 means all show that these two vowels are lowered in pre-velar context. Furthermore, Post-Hoc tests show that the allophonic distinction is maintained across all speaker groups. This is consistent with previous descriptions of these two vowels in Hong Kong Cantonese (Yue-Hashimoto, 1972). There is hence continuity in at least this part of the vowel system of HCAN. This constitutes the evidence that HL speakers maintain not only phonological contrasts, but also allophonic conditioning.

7.2 External factors

Having established that allophonic conditioning in pre-velar context is maintained in Toronto HCAN, the next question to address is whether or not there is evidence for variation in the allophones of these vowels based on generational group or sex (external factors). The results show that these two external factors are significant only if they are grouped together as a single factor group in terms of accounting for inter-speaker variation in the pronunciation of the allophones of /iː/. Neither generational group nor sex showed significance in any of the tests in which they were treated as separate factors in accounting for variation in /iː/ as a whole. A close examination of the results shows why this may be the case.

If we go back to the results of the Post-Hoc Tests for [i] shown in **Table** 0, we observe that GEN 2 female speakers are significantly different from all other speaker groups in F2 production including GEN 1 female speakers. In fact, the GEN 2 female – GEN 1 female pair was also the only pair that showed a significant difference in the F2 of [ɪk/ɪŋ] (see **Table 2**1). This suggests that sex-based patterns (and possibly gender-based patterns) show inter-generational differences. Similarly, if we return to the comparison of the vowel spaces for the four speaker groups (Figures 3–6), we observed different patterns of vowel overlap based on generational group. GEN 1 male speakers showed some overlap between [ɪk/ɪŋ] and /ɛ/ while GEN 1 female speakers showed no overlap based on mean F1/F2 ± 1 standard deviation. Among GEN 2 speakers, the sex based pattern is reversed with GEN 2 female speakers showing more overlap than GEN 2 male speakers.

Thus, the lack of significant effects for generation and sex as distinct factors is due to an interaction between the two factors. One major generational difference is that sex based patterns are different for each generational group reflecting a change in socially conditioned patterns of

variation. In some cases, sex based patterns may even be completely reversed. Kang and Nagy (2012) report similar findings in a comparison between Homeland and Toronto Heritage Korean. They found that while VOT merger is a recent female led sound change in Homeland Korean, the GEN 2 HL speakers showed a lack of sex based differences. Both young male and female GEN 2 HL Korean speakers participate in the change.

One explanation for the inter-generational change in sex based patterns could be that traditional gender roles in Hong Kong are different from gender roles in contemporary Canadian society. Being male or female in Hong Kong may mean something very different from being male or female in Canada. This may possibly motivate inter-generational differences in sociophonetic variation. This explanation would however require further research.

Another explanation is that these patterns are simply an artifact of the data analyzed. This explanation is supported by the considerable amount of individual variation in both GEN 1 and GEN 2 speakers discussed in Section 6.2.1. The r^2 value for random effects in the model shown in **Table** 8 is almost twice as high as the r^2 value for fixed effects. This means that individual speaker and word account for a larger percentage of variation in the data analyzed than generational group and sex (whether or not considered as distinct variables). GEN 1 speakers exhibit variation as do GEN 2 speakers. GEN 2 HCAN speakers are not uniform in their phonetic production of allophones and neither are GEN 1 speakers. Sex may be one factor based on the results of the present study, but there also may be other overlapping factors that have not yet been considered. Thus, before we can address the inter-generational change in sex based patterns, we need to confirm its presence with a larger sample of speakers.

7.3 Accounting for variation

Having established that Toronto HCAN speakers vary in their phonetic production of the allophones of /iː/, the more challenging problem is accounting for the observed variation. What is the motivation behind inter-generational change (assuming the same patterns hold if more speakers are analyzed)? Is it due to internal motivation, or due to phonetic assimilation with the most phonetically similar TOENG vowels, or due to phonological influence from TOENG? The results show the strongest support for phonological influence. Yet at the same time, the two other accounts cannot be completely ruled out until additional comparative studies are completed as advocated by Nagy (2011).

According to the internal motivation hypothesis, there may be influence based on pre-existing variation among GEN 1 speakers that originated in Hong Kong Cantonese. The results from this study show that there is indeed inter-speaker variation among GEN 1 speakers. What is less conclusive is how pre-existing variation influences or does not influence the pronunciation of allophones among GEN 2 speakers. If there is influence, this hypothesis predicts that there would be additional movement of [ɪk/ɪŋ] towards [ek/eŋ] and [ʊk/ʊŋ] towards [ok/oŋ]. This would result in higher F1 (lowering) for [ɪk/ɪŋ] and [ʊk/ʊŋ] and possibly lower F2 (retraction) for [ɪk/ɪŋ]. The results show evidence of change only in the direction of lower F2 for [ɪk/ɪŋ]. We do not see evidence of additional vowel lowering (higher F1) advanced by GEN 2 speakers. Instead, we see only evidence of retraction with GEN 2 female speakers having the most retracted variants of [ik/ɪŋ]. If there already was retraction in Hong Kong Cantonese, the retraction among GEN 2 female speakers could be an internally motivated change following the same trajectory of change initiated in Hong Kong.

The problem with supporting this hypothesis is the lack of existing sociophonetic research on Hong Kong Cantonese vowels. Without comparable data from Hong Kong, it is uncertain whether the primary direction of change from [ɪk/ɪŋ] to [ek/eŋ] is in F1 or in F2. The lack of significant inter-group differences in F1 for both vowels also suggests that the lowering may already have been a nearly completed change in Hong Kong Cantonese. Results also show [iː] fronting among GEN 2 female speakers, a change that was not previously expected. Interestingly, Lee (1983) observed that Hong Kong [iː] is more peripheral than Guangzhou [iː]. It could be possible that this is a recent sound change in Hong Kong that did not affect GEN 1 Toronto speakers. If this is the case, it could be possible that the peripheralization of this vowel among some GEN 2 speakers is internally motivated. Yet, the problem is lack of apparent time data from Hong Kong. A future cross-community comparison with Hong Kong Cantonese as advocated by Nagy (2011) would be essential in determining whether or not these changes can be attributed to internal motivation. If these changes do not occur in Hong Kong, then that would weaken support for internal motivation and strengthen support for the second and third hypotheses, which are both based on contact with TOENG.

According to the second hypothesis, the [ɪk/ɪŋ] retraction is influenced by phonetic assimilation with TOENG. If TOENG /ɪ/ is retracted (Roeder & Jarmasz, 2010), then the retraction observed among GEN 2 female speakers could be due to assimilation with the most phonetically similar vowel in TOENG. This hypothesis may also account for lack of [uː] and [ʊk/ʊŋ] fronting among GEN 2 speakers. Perhaps [uː] and [ʊk/ʊŋ] do not front in HCAN because they are not phonetically similar enough to TOENG /u/ and /ʊ/. This is exactly what Godson (2004) suggested for the lack of /u/ fronting in HL western Armenian. Interestingly, Godson (2004) also argued that phonetic similarity influenced only the front vowels in HL

western Armenian because they have more similar counterparts in the local dialect of English. This is quite similar to what was observed in the present HCAN study. Evidence of change was observed only for front vowels. Thus, perhaps the retraction of [ɪk/ɪŋ] is due to assimilation with TOENG [ɪk/ɪŋ] because this part of the HCAN vowel system shares more similarities. Phonetic assimilation could also explain the [iː] fronting among GEN 2 female speakers, if it turns out that TOENG /i/ is more fronted than HCAN [iː].

Existing studies of TOENG vowels have focused largely on ongoing change rather than on cross-linguistic comparison and have also produced conflicting results about trajectories of change. This makes it difficult to assess whether or not phonetic assimilation with TOENG is actually happening. Roeder and Jarmasz (2010), for example, show that retraction is the primary direction of change in TOENG /ɪ/. This supports the phonetic assimilation hypothesis since retraction is the same direction of movement observed in the HCAN data. The speakers analyzed in this study, however, did not include Torontonians who are bilingual in English and Cantonese unlike Hoffman (2010)[12]. Yet, the results of Hoffman (2010) do not provide supporting evidence for phonetic assimilation. They show an overall lack of ethnic based distinctions. They also show that sex and age are significant factors with young male speakers leading in the lowering of TOENG /ɪ/. If this is true, the phonetic assimilation hypothesis would predict that GEN 2 male speakers would lead in the lowering of [ɪk/ɪŋ]. Instead, we observe GEN 2 female speakers leading in the retraction of [ɪk/ɪŋ]. To be conclusively sure about whether or not phonetic assimilation with TOENG is a valid account of observed variation in the HCAN data, a cross-linguistic comparison (Nagy, 2011) would be needed that would involve normalizing a complete set of data of TOENG vowels together with Toronto HCAN vowels. Would such an analysis show a pattern of increasing overlap with TOENG vowels among GEN 2 HCAN speakers?

Finally, the third hypothesis, which is based on phonological influence, can also account for the retraction of [ɪk/ɪŋ] as well as the fronting of [iː] among GEN 2 female speakers. If HL speakers maintain more of a cross-linguistic phonetic distinction between similar sounds due to their early exposure to two different languages as Chang et al. (2011) have shown, then this means that HL speakers have a combined phonological inventory that is larger than the single phonological inventory of monolingual speakers. This could make it more likely for HL speakers to recognize fine-grained phonetic distinctions. If this is true, then this would explain how some of the GEN 2 speakers may be reinterpreting HCAN allophones as phonologically distinct units under influence from a language in which similar phones are different phonemes. If allophones in the HL are recognized as distinct phonological units, changes in their pronunciation may be influenced by an attempt to maintain cross-linguistic distinctions. This would account for both the fronting of [iː] and the retraction of [ɪk/ɪŋ] among GEN 2 females. The outcome of these two directions of movement is increasing phonetic difference between these two allophones. This suggests that GEN 2 female speakers may be reinterpreting the allophonic distinction as a phonologically contrastive one within the Cantonese system while at the same time increasing the difference between these two allophones as a way of maintaining cross-linguistic distinctions between their Cantonese vowels and TOENG vowels.

The phonological influence hypothesis may also account for the lack of change in the allophones of /uː/. If both [uː] and [ʊ] are already phonetically distinct from TOENG /u/, perhaps they remain stable as a way of maintaining cross-linguistic distinctions with their closest TOENG counterparts. The lack of [uː] and [ʊk/ʊŋ] fronting is exactly the same pattern found in studies of HL Mandarin and HL western Armenian in which a high back vowel does not front under influence from /u/-fronting in English. Like Mandarin, Cantonese also has a phonological

contrast between two high round vowels (/u/ vs. /y/), which could further contribute to a lack of /u/-fronting. The lack of /u/-fronting across three HLs under contact with English suggests that this may be a broader pattern found across HL contact situations. This is evidence of maintenance for this part of the Cantonese vowel system among HL speakers.

What would strengthen the phonological influence hypothesis is a comparative study that includes HCAN vowels normalized together with TOENG vowels from bilingual Torontonians. Increasing distinction would support the phonological influence hypothesis while decreasing distinction would support phonetic assimilation.

8. Conclusion

This paper has shown maintenance of allophonic conditioning of /iː/ and /uː/ across two generations of speakers of Toronto HCAN. This paper has also shown evidence for sociophonetic variation in the pronunciation of allophones [iː] and [ɪk/ɪŋ] with GEN 2 female speakers leading in both the fronting of [iː] and in the retraction of [ɪk/ɪŋ]. This paper discussed three possible accounts for the observed patterns based on internal motivation, phonetic assimilation, and phonological influence. In assessing each account, this paper discussed how Nagy's (2011) three proposed set of comparisons would resolve the uncertainty about the best explanation.

HL phonology is still a relatively new field of research interest. Previous research has shown maintenance of phonological contrasts in western Armenian (Godson, 2004), Mandarin (Chang et al., 2011), and Spanish (Ronquest, 2013). The contribution of this paper is in showing that allophonic conditioning is also maintained, at least among Toronto HCAN speakers. Furthermore, this study also shows inter-speaker variation among both GEN 1 and GEN 2

speakers. Future research needs to consider inter-speaker variation among GEN 1 speakers in addition to influence from the dominant language, which could either mean phonetic assimilation (as is widely observed in L2 phonology) or phonological influence. The possibility of phonological influence is one that should especially be investigated since it may be a unique characteristic of HL contact situations. This uniqueness may be facilitated by community-wide early exposure to two distinct languages with distinct vowel systems. Chang et al. (2011) is perhaps the only study that has explicitly discussed this possibility. The results from this paper show evidence that this may be possible among HCAN speakers.

This paper also contributes to the Cantonese linguistics literature by introducing a variationist approach to the study of Cantonese vowels. The study presented in this paper is the first sociophonetic study of vowels in Cantonese using large-scale corpus data. It examined a much larger sample of Cantonese speakers (N = 20) than Lee's (1983) study (N = 6). As discussed in Section 2.3, the transcription of high vowel allophones has been the subject of debate among Cantonese linguists. This paper contributes to this literature by showing evidence for inter-speaker variation among both GEN 1 and GEN 2 speakers. Inter-speaker variation may account for previous discrepancies in the transcription of these vowels.

Finally, this paper lays out the basis for future studies of HCAN vowels as well as studies of vowels in other HLs. Although this paper is less conclusive about how to account for the variation observed, it is conclusive in showing multiple possible sources of variation in HLs. This study discussed three possibilities and showed how three different explanations (internal motivation, phonetic assimilation, and phonological influence) can account for the [iː] fronting and [ɪk/ɪŋ] retraction observed among GEN 2 female speakers. To further assess these three accounts, future work needs to address whether or not the same changes are found in Hong Kong

Cantonese and how the TOENG vowels of English-Cantonese bilingual speakers compare to their Cantonese vowels. A cross-community comparison showing the same changes in Hong Kong Cantonese would support the internal motivation hypothesis. A cross-linguistic comparison with TOENG showing increasing similarity would support phonetic assimilation while decreasing similarity would support phonological influence. Future research also needs to address whether or not the generation and sex interaction observed holds for a larger group of speakers. If this pattern does hold, the next question to ask would be why there are different patterns based on generational group. To what extent could this be related to a change in the societal setting in which Cantonese is spoken?

HL phonology remains a largely underexplored area. By avoiding a deficit perspective and by using the tools of variationist sociolingusitics to examine the vowel system of a HL, this paper has shown how HLs can contribute to the development of sociolinguistics and has set a useful foundation for future work on HL vowels. In spite of the limitations of this study, the results presented show many opportunities for future work on HCAN and more generally on HLs within a variationist framework. It also underscores the importance of investigating all languages spoken within a particular multilingual community.

Acknowledgements

This paper originally developed from a comprehensive paper written in partial fulfillment of the degree requirements for a PhD degree at the University of Pittsburgh (http://d-scholarship.pitt.edu/20436/). The HLVC Corpus is funded by a research grant from the Social Sciences and Humanities Research Council of Canada 410-2009-2330 (2009-2012) and is developed by Naomi Nagy, Alexei Kochetov, Yoonjung Kang, and James Walker. The author

would like to thank Naomi Nagy for access to the corpus as well as for overall guidance and support, Scott Kiesling and Shelome Gooden for feedback on earlier versions of this paper, Rachel Coulter for copy editing assistance, Yutian Tan for translating the abstract, audience feedback from presentations at the University of Pittsburgh, University of Toronto (NWAV 44), and Memorial University of Newfoundland (APLA 39 ALPA), and all the HLVC Research Assistants (full list at: http://projects.chass.utoronto.ca/ngn/HLVC/0_0_home.php).

References

Bauer, Robert S., & Benedict, Paul K. (1997). *Modern Cantonese phonology*. Berlin: Mouton de Gruyter.

Bauer, Robert S., Cheung, Kwan-hin., & Cheung, Pak-man (2003). Variation and merger of the rising tones in Hong Kong Cantonese. *Language Variation and Change*, 15(2), 211–225.

Boersma, Paul, & Weenink, David (2014). Praat: Doing phonetics by computer (Version 5.4.02) [Computer program]. Retrieved December 14, 2014 from http://www.praat.org/

Chang, Charles B., Yao, Yao, Haynes, Erin F., & Rhodes, Russell (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America*, 129(6), 3964–3980. DOI: 10.1121/1.3569736

Clarke, Sandra, Elms, Ford, & Youssef, Amani (1995). The third dialect of English: Some Canadian evidence. *Language Variation and Change*, 7(2), 209–228. DOI: 10.1017/S0954394500000995

Godson, Linda (2004). Vowel production in the speech of western Armenian heritage speakers. *Heritage Language Journal*, 2(1), 1–26.

Hoffman, Michol F. (2010). The role of social factors in the Canadian vowel shift: Evidence from Toronto. *American Speech*, 85(2), 121–140. DOI: 10.1215/00031283-2010-007

Hoffman, Michol F., & Walker, James A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, 22(1), 37–67. DOI: 10.1017/S0954394509990238

Hsu, Madeline Y. (2000). *Dreaming of gold, dreaming of home: Transnationalism and migration between the United States and South China, 1882-1943*. Stanford: Stanford University Press.

Johnson, Daniel E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, 3(1), 359–383.

Kang, Yoonjung, & Nagy, Naomi (2012). VOT merger in heritage Korean in Toronto. *Proceedings of the Annual Meeting of the Canadian Linguistic Association*. Available fromhttp://www.yoonjungkang.com/uploads/1/1/6/2/11625099/kangnagy_2012_vot_cla.pdf

Labov, William (1994). *Principles of linguistic change, volume 1*: *Internal factors*. Oxford, UK; Cambridge, MA: Blackwell.

Labov, William, Ash, Sharon, & Boberg, Charles (2006). *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.

Lau, Chun-fat (2003). Labovian principles of vowel shifting revisited: The short vowel shift in New Zealand English and Southern Chinese. In Barry J. Blake & Kate Burridge (Eds.), *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13-17 August 2001* (pp. 293–301). Amsterdam: John Benjamins. DOI: 10.1075/cilt.237

Lee, Thomas (1983). The vowel system in two varieties of Cantonese. *UCLA Working Papers in Phonetics*, 57, 97–114.

Li, Peter S. (2005). The rise and fall of Chinese immigration to Canada: Newcomers from Hong Kong Special Administrative Region of China and mainland China, 1980-2000. *International Migration*, 43(3), 9–32.

Man, Evelyn Y. (2006). First language use and language behavior of Chinese students in Toronto, Canada. In Kimi Kondo-Brown (Ed.), *Heritage language development: Focus on East Asian immigrants* (pp. 209-241). Amsterdam: John Benjamins.

Matthews, Stephen, & Yip, Virginia (2011). *Cantonese: A comprehensive grammar*. London: Routledge.

Mok, Peggy K., Zuo, Donghui, & Wong, Peggy 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.

Montrul, Silvina A. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.

Nagy, Naomi (2011). A multilingual corpus to explore variation in language contact situations. *Rassegna Italiana Di Linguistica Applicata*, 43(1/2), 65–84.

Polinsky, Maria (2011). Heritage languages. In Mark Aronoff (Ed.), *Oxford bibliographies in linguistics* [online resource]. DOI: 10.1093/obo/9780199772810-0067

Polinsky, Maria, & Kagan, Olga (2007). Heritage languages: In the "Wild" and in the classroom. *Language and Linguistics Compass*, 1(5), 368–395. DOI: 10.1111/j.1749-818X.2007.00022.x

Roeder, Rebecca, & Jarmasz, Lidia-Gabriela (2010). The Canadian shift in Toronto. *The Canadian Journal of Linguistics*, 55(3), 387–404. DOI: 10.1353/cjl.2010.0013

Ronquest, Rebecca E. (2013). An acoustic examination of unstressed vowel reduction in Heritage Spanish. In Chad Howe, Sarah E. Blackwell, & Margaret Lubbers Quesada (Eds.), *Selected proceedings of the 15th Hispanic Linguistics Symposium* (pp. 157–171). Available from http://www.lingref.com/cpp/hls/15/paper2882.pdf

Siemiatycki, Myer, Rees, Tim, Ng, Roxana, & Rahi, Khan (2003). Integrating community diversity in Toronto: On whose terms? In Paul Anisef & C. Michael Lanphier (Eds.), *The world in a city* (pp. 373–456). Toronto: University of Toronto Press.

Sloetjes, Han, & Wittenburg, Peter (2008). Annotation by category – ELAN and ISO DCR. In *Proceedings of the sixth International Conference on Language Resources and Evaluation (LREC 2008)*, 816-820. Available from http://www.lrec-conf.org/proceedings/lrec2008/

Statistics Canada (2012). Visual Census 2011, Ottawa. Retrieved January 1, 2015, from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/vc-rv/index.cfm?Lang=eng

Stevens, Kenneth N., & House, Arthur S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech & Hearing Research*, 6(2), 111–128.

Taylor, Bill (2007, December 30). The language quilt: From irregular stitching, an ongoing work of art. *The Toronto Star*, p. ID06

Thomas, Erik, & Kendall, Tyler (2007). NORM: The vowel normalization and plotting suite (Version 1.1) [Computer program]. Retrieved December 14, 2014 from http://ncslaap.lib.ncsu.edu/tools/norm/

Thompson, Richard H. (1989). *Toronto's Chinatown: The changing social organization of an ethnic community*. New York: AMS Press.

Yue-Hashimoto, Oi-kan (1972). *Phonology of Cantonese*. Cambridge, England: Cambridge University Press.

Zee, Eric (2003). Frequency analysis of the vowels in Cantonese from 50 male and 50 female speakers. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1117–1120). Available from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_1117.pdf

Endnotes

[1] This number is up to date as of February 1, 2016. The eight languages include Cantonese, Faetar, Hungarian, Italian, Korean, Polish, Russian, and Ukrainian.

[2] From this point on, I will use rime group format interchangeably with allophones. The Cantonese rime groups discussed in this paper are [ɪk/ɪŋ] and [ʊk/ʊŋ]. The rime group is a unit of phonological analysis that is widely used among Chinese linguists. It consists of a vowel nucleus followed by an optional coda consonant. The advantage of using the rime group format in this paper is in making it easier to distinguish between English and Cantonese vowels.

[3] I am uncertain about why vowel lowering occurs only before velar consonants. Some research has shown that velar consonants have a different effect from what is observed for Cantonese. Stevens and House (1963), for example, show that velar consonants have the least amount of co-articulation effects on adjacent vowels compared to labial and alveolar consonants.

In Cantonese, /iː/ can be followed by both labial and alveolar consonants but in both cases, the vowel is realized as [iː] rather than as [ɪ] (Yue-Hashimoto, 1972). As far as I can tell, the lowering effect of velar consonants in Cantonese is a language-specific phenomenon.

[4] The 20 speakers analyzed in this paper are a subset of the 40 speakers in the complete corpus. A long-term goal is to have all 40 speakers analyzed. For this initial study, criteria for selecting speakers included audio quality (a few sound files were corrupt and not usable), transcription completeness (some files still need to be processed), and talkativeness (more talkative speakers were selected since these speakers would likely have more usable vowel tokens). An attempt was also made to balance the number of speakers based on generation, sex, and age. The corpus also includes GEN 3 speakers, defined as those born in the GTA and that also have parents that qualify as GEN 2 speakers. GEN 3 speakers were not included since this paper focuses on differences between GEN 1 and GEN 2 speakers.

[5] The effect of exposure to English in Hong Kong on the Cantonese of GEN 1 speakers is worth investigating in future research. This would require a different analysis that is beyond the scope of this paper.

[6] The guidelines for the HerLD Corpus allow speakers who moved to the GTA up to the age of six to qualify as a GEN 2 speaker. The specific speakers analyzed in this paper, however, have all lived in the GTA since the age of three or earlier.

[7] Languages other than English and Cantonese were also reported by both GEN 1 and GEN 2 speakers. This includes French as well as other varieties of Chinese including Mandarin, Siyi dialects especially Taishanese/Toisanese, and the Chaozhou dialect. Knowledge of other

languages and dialects, however, was variable and not systematically recorded making it difficult to assess the extent to which such knowledge may have an effect on the results of this study.

[8] This is similar to the number of vowel tokens per speaker collected as part of the Atlas of North American English (Labov, Ash, & Boberg, 2006).

[9] Tokens of [uːn] were also included for speakers who had less than five tokens of [uː] in open syllable context. What is important here is that even with a coda [n], /uː/ is realized as [uː] rather than as [ʊ].

[10] Although different settings are typically used for male and female speakers to address the problem of different vocal tract lengths, the procedure adopted here takes a different approach to addressing this problem by eye checking for fit between the dark formant bands on the spectrogram and the Praat formant tracker calculations on an individual token basis. Ultimately, it is accurate measurements of individual tokens that matter rather than whether a speaker is male or female per se because there can be quite a bit of variation in vocal tract length among both males and females.

[11] Tone was also included for /uː/ since this was the only vowel with tokens from different tone categories.

[12] Hoffman and Walker (2010) also found lack of evidence for the emergence of "ethnolects" although they do discuss the possibility that some individual speakers may be using features such as /t/-deletion to index ethnic identity more so than others.

TABLES

Table 1. Cantonese monophthongs

Front		Central	Back
Unrounded	Rounded		
i	y		u
ɛ	œ	ɐ	ɔ
		a	

Table 2. Changes predicted for each allophone

		Hypotheses		
Vowel Category	Allophone	Internal motivation	Phonetic assimilation	Phonological influence
/iː/	[iː]	--	--	Fronting and/or raising
	[ɪk/ɪŋ]	Lowering, with some retraction	Retraction, with some lowering (Canadian Shift)	--
/uː/	[uː]	--	Fronting	Raising and/or backing
	[ʊk/ʊŋ]	Lowering	Fronting	--

Table 3. List of speakers from the HerLD corpus (Nagy, 2011)

	Male	Female	Total
GEN 1	C1M46A C1M52A C1M59A C1M61A C1M62A	C1F50A C1F54A C1F58A C1F78A C1F82A	= 10
GEN 2	C2M16A C2M17B C2M21D C2M27A C2M44A	C2F16A C2F16B C2F16C C2F20A C2F21B	= 10
	= 10	= 10	TOTAL N = 20 speakers

Note. The speaker codes indicate the following: "C" = Cantonese, "1" or "2" = Generation Group, "M" or "F" = Male or Female, Two-digit number indicates age, Last character is used to distinguish between multiple speakers with same demographic characteristics (i.e., A, B, C, etc.)

Table 4. Token distribution for each speaker

Vowel	Open syllable	Pre-velar	Total
/aː/	15 [aː]	0	= 15
/ɛː/	10 [ɛː]	5 [ɛːk]/[ɛːŋ]	= 15
/iː/	10 [iː]	5 [ɪk]/[ɪŋ]	= 15
/ɔː/	10 [ɔː]	5 [ɔːk]/[ɔːŋ]	= 15
/uː/	5 [uː][a]	10 [ʊk]/[ʊŋ]	= 15
	= 50	= 25	TOTAL N = 75 per speaker

[a]Not all tokens of [uː] occur in open syllable context.

Table 5. Independent variables examined for both F1 and F2

Random Effects	Fixed Effects		
	Social Factors	Linguistic Factors	Factor Group
Speaker, Word	Generation, Sex	Following Velar, Tone	GEN×Sex×Velar

Table 6. Best step-down model for F1 of /iː/

Random Effects (r^2 = 0.079)				
Speaker, Word				
Fixed effects (r^2 = 0.27)				
	Coef.	N	Mean (Hz)	*p*-value
Velar				< .001***
Yes	26.23	100	427	
No	-26.23	100	373	
Generation				n.s.
Sex				n.s.

Note. r^2 [total] = 0.349

Table 7. Best step-down model for F1 of /iː/ with GEN×Sex×Velar factor group included

Random effects (r^2 = 0.076)				
Speaker, Word				
Fixed Effects (r^2 = 0.275)	**Coef.**	**N**	**Mean (Hz)**	***p*-value**
Gen×Sex×Velar				< .01**
2F [ɪk/ɪŋ]	34.908	25	440	
1M [ɪk/ɪŋ]	27.600	25	428	
1F [ɪk/ɪŋ]	22.644	25	417	
2M [ɪk/ɪŋ]	18.749	25	423	
1F [iː]	-22.336	50	377	
2M [iː]	-24.120	50	374	
2F [iː]	-24.651	50	375	
1M [iː]	-32.794	50	366	

Note. r^2 [total] = 0.351

Table 8. Best step-down model for F2 of /iː/ with GEN×Sex×Velar factor group included

Random effects (r^2 = 0.21)				
Speaker, Word				
Fixed effects (r^2 = 0.126)				***p*-value**
	Coef.	N	Mean (Hz)	
GEN×Sex×Velar				< .001***
2F [iː]	111.817	50	2017	
1M [iː]	25.805	50	1929	
2M [iː]	10.632	50	1911	
1F [iː]	-10.183	50	1890	
2M [ɪk/ɪŋ]	-14.341	25	1819	
1F [ɪk/ɪŋ]	-20.962	25	1894	
1M [ɪk/ɪŋ]	-29.989	25	1863	
2F [ɪk/ɪŋ]	-72.779	25	1781	

Note. r^2 [total] = 0.336

Table 9. Tukey post-hoc tests of difference between [iː] and [ɪk/ɪŋ]

Pair 1	Pair 2	F1	F2
G1F [iː]	G1F [ɪk/ɪŋ]	**	n.s.
G1M [iː]	G1M [ɪk/ɪŋ]	***	n.s.
G2F [iː]	G2F [ɪk/ɪŋ]	***	***
G2M [iː]	G2M [ɪk/ɪŋ]	***	***

Note. **p < .01, ***p < .001

Table 10. Tukey post-hoc tests of inter-group differences in [i:]

Pair 1	Pair 2	F1	F2
G2F	G1M	n.s.	*
G2F	G1F	n.s.	***
G2F	G2M	n.s.	**
G2M	G1M	n.s.	n.s.
G2M	G1F	n.s.	n.s.
G1F	G1M	n.s.	n.s.

Note. *p < .05, **p < .01, ***p < .001

Table 21. Tukey post-hoc tests of inter-group differences in [ɪk/ɪŋ]

Pair 1	Pair 2	F1	F2
G2F	G1M	n.s.	n.s.
G2F	G1F	n.s.	*
G2F	G2M	n.s.	n.s.
G2M	G1M	n.s.	n.s.
G2M	G1F	n.s.	n.s.
G1F	G1M	n.s.	n.s.

Note. *p < .05

Table 32. Best step-down model for F1 of /u:/ (r^2 [total] = 0.26)

Random effects (r^2 = 0.033)				
Speaker, Word				
Fixed effects (r^2 = 0.227)				***p*-value**
	Coef.	N	Mean (Hz)	
Velar				< .001***
Yes	26.48	202	440	
No	-26.48	98	388	
Generation				n.s.
Sex				n.s.
Tone				n.s.

Note. r^2 [total] = 0.26

Table 43. Best step-down model for F1 of /uː/ with GEN×Sex×Velar factor group included

Random effects (r^2 = 0.044)				
Speaker, Word				
Fixed effects (r^2 = 0.238)	Coef.	N	Mean (Hz)	*p*-value
GEN×Sex×Velar				< .001***
1M [ʊk/ʊŋ]	34.070	47	447	
2M [ʊk/ʊŋ]	27.602	50	441	
1F [ʊk/ʊŋ]	23.135	49	437	
2F [ʊk/ʊŋ]	23.101	56	436	
2M [uː]	-16.686	25	397	
1F [uː]	-17.061	26	396	
1M [uː]	-29.361	28	385	
2F [uː]	-44.799	19	371	
Tone				n.s.

Note. r^2 [total] = 0.282

Table 54. Best step-down model for F2 of /uː/ with GEN×Sex×Velar factor group included (r^2 [total] = 0.357)

Random effects (r^2 = 0.085)				
Speaker, Word				
Fixed effects (r^2 = 0.272)	Coef.	N	Mean (Hz)	*p*-value
GEN×Sex×Velar				< .05*
2M [uː]	71.711	25	1236	
1M [uː]	63.067	28	1238	
1F [uː]	18.771	26	1185	
2F [uː]	13.287	19	1185	
2M [ʊk/ʊŋ]	-1.642	50	1277	
1F [ʊk/ʊŋ]	-49.100	49	1243	
1M [ʊk/ʊŋ]	-52.586	47	1234	
2F [ʊk/ʊŋ]	-63.508	56	1236	
Tone				n.s.

Note. r^2 [total] = 0.357

Table 65. Post-hoc results summary

Vowel	Formant	Allophones	Overall Results
/i/	F1	[iː] [ɪk/ɪŋ]	Pre-velar lowering for all groups
	F2	[iː] [ɪk/ɪŋ]	GEN 2 Females > All other groups GEN 2 Females < GEN 1 Females
/u/	F1	[uː] [ʊk/ʊŋ]	Pre-velar lowering for all groups
	F2	[uː] [ʊk/ʊŋ]	No significant inter-group differences No significant inter-group differences

FIGURES

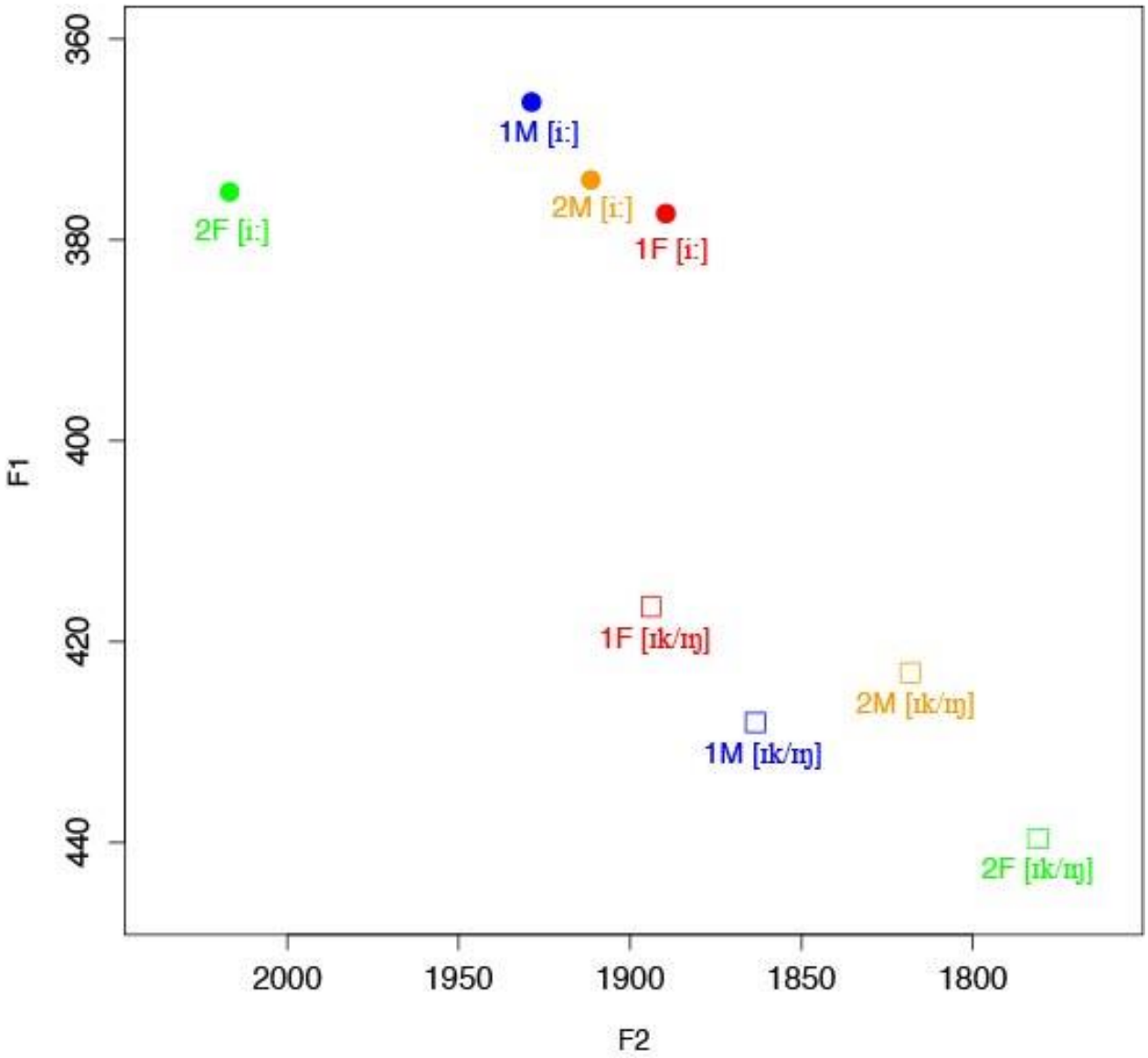

Figure 1. Plot of mean F1/F2 of /i:/ for each speaker group

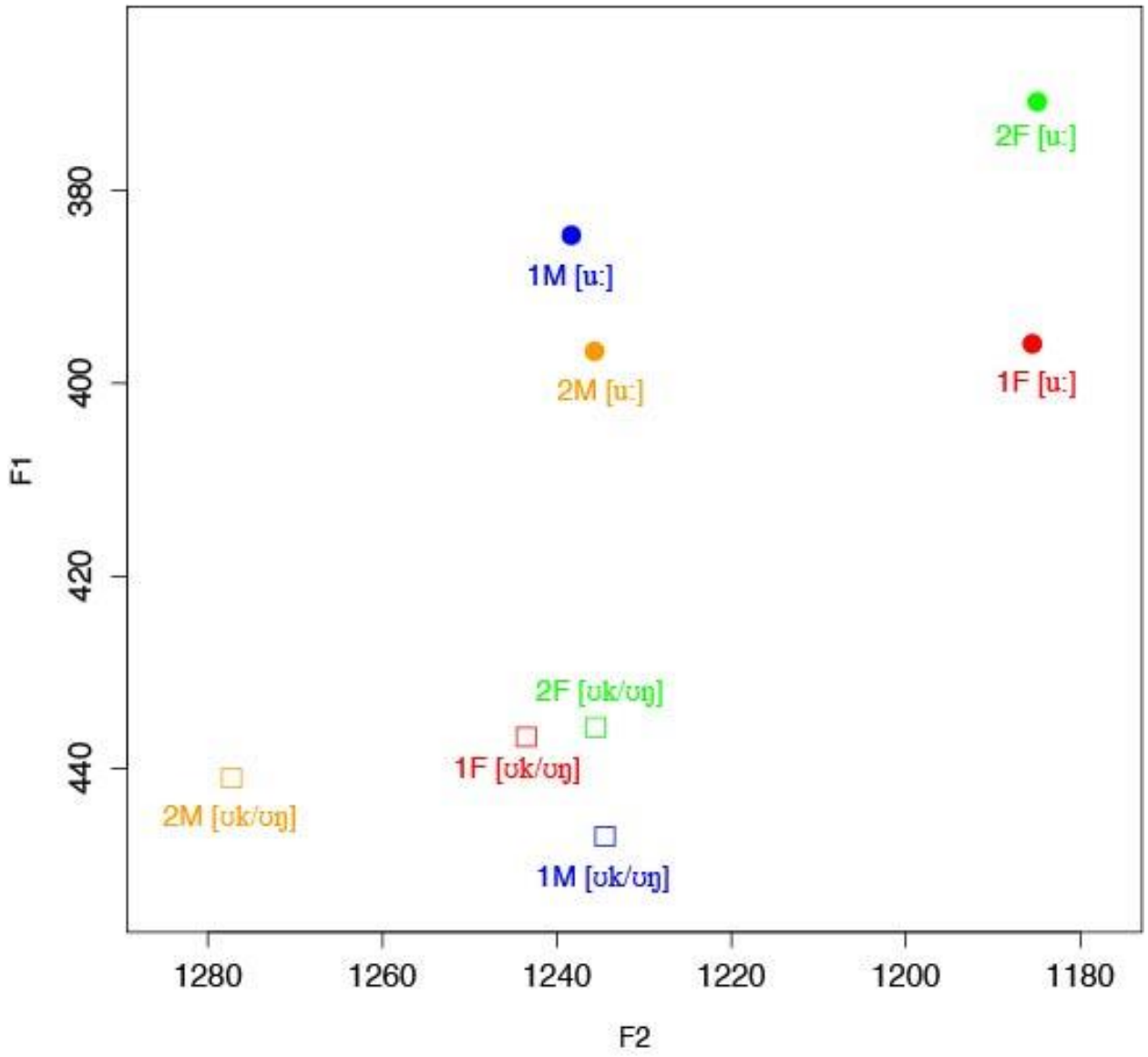

Figure 2. Plot of mean F1/F2 of /uː/ for each speaker group

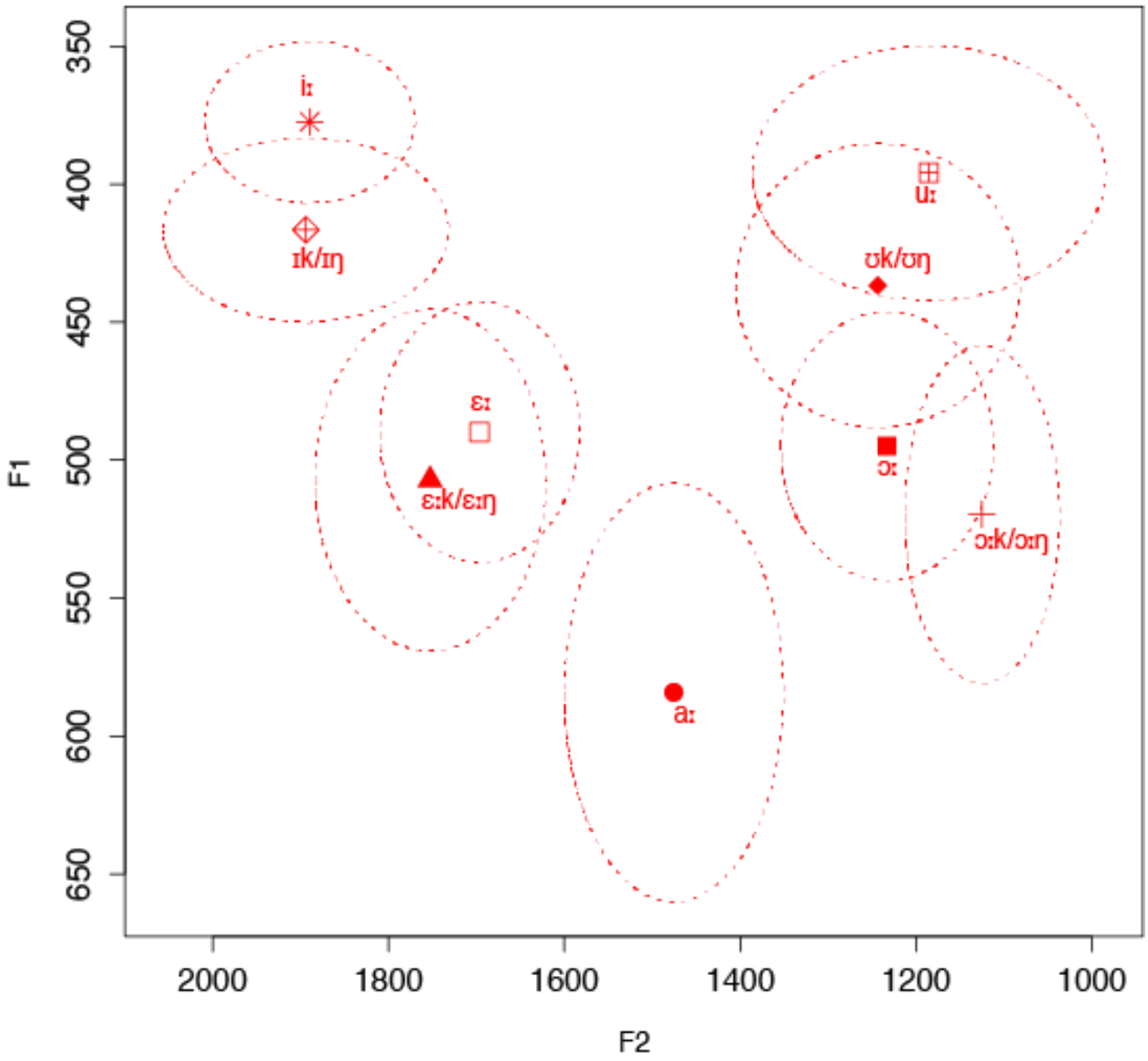

Figure 3. GEN 1 female speakers, Mean F1/F2 with ± 1 standard deviation, Normalized using the Lobanov method.

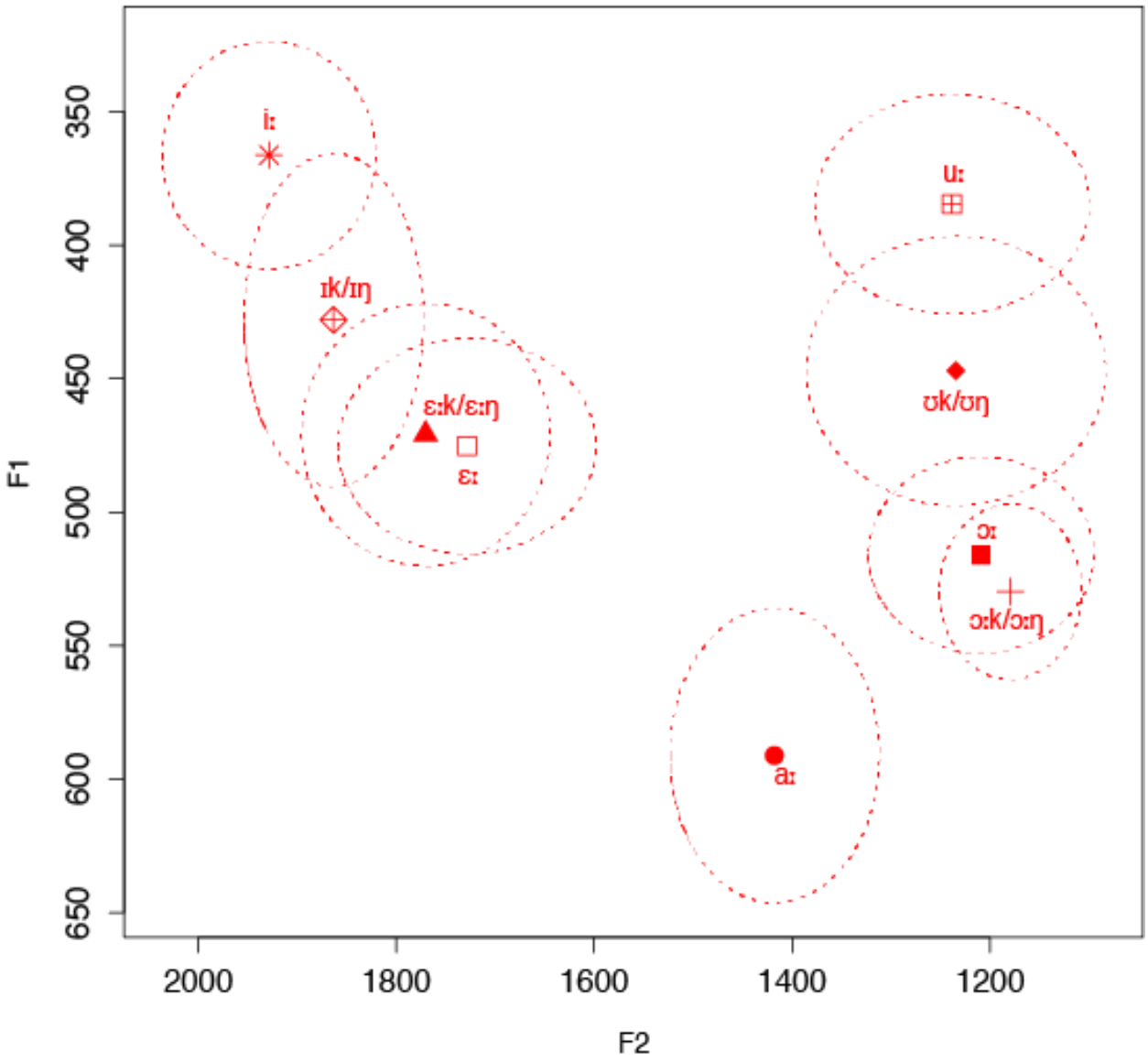

Figure 4. GEN 1 male Speakers, Mean F1/F2 with ± 1 standard deviation, Normalized using the Lobanov method.

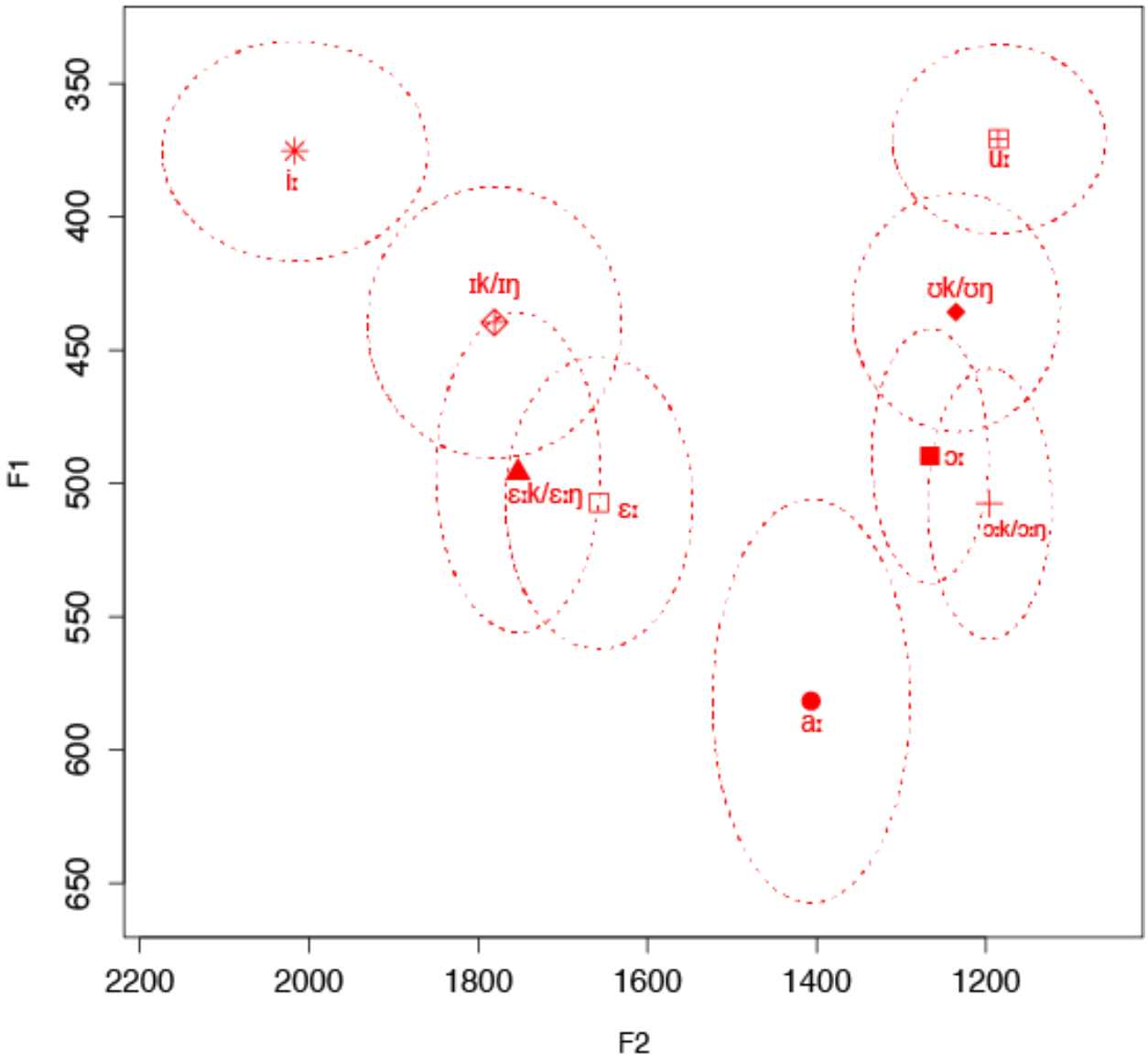

Figure 5. GEN 2 female speakers, Mean F1/F2 with ± 1 standard deviation, Normalized using the Lobanov method.

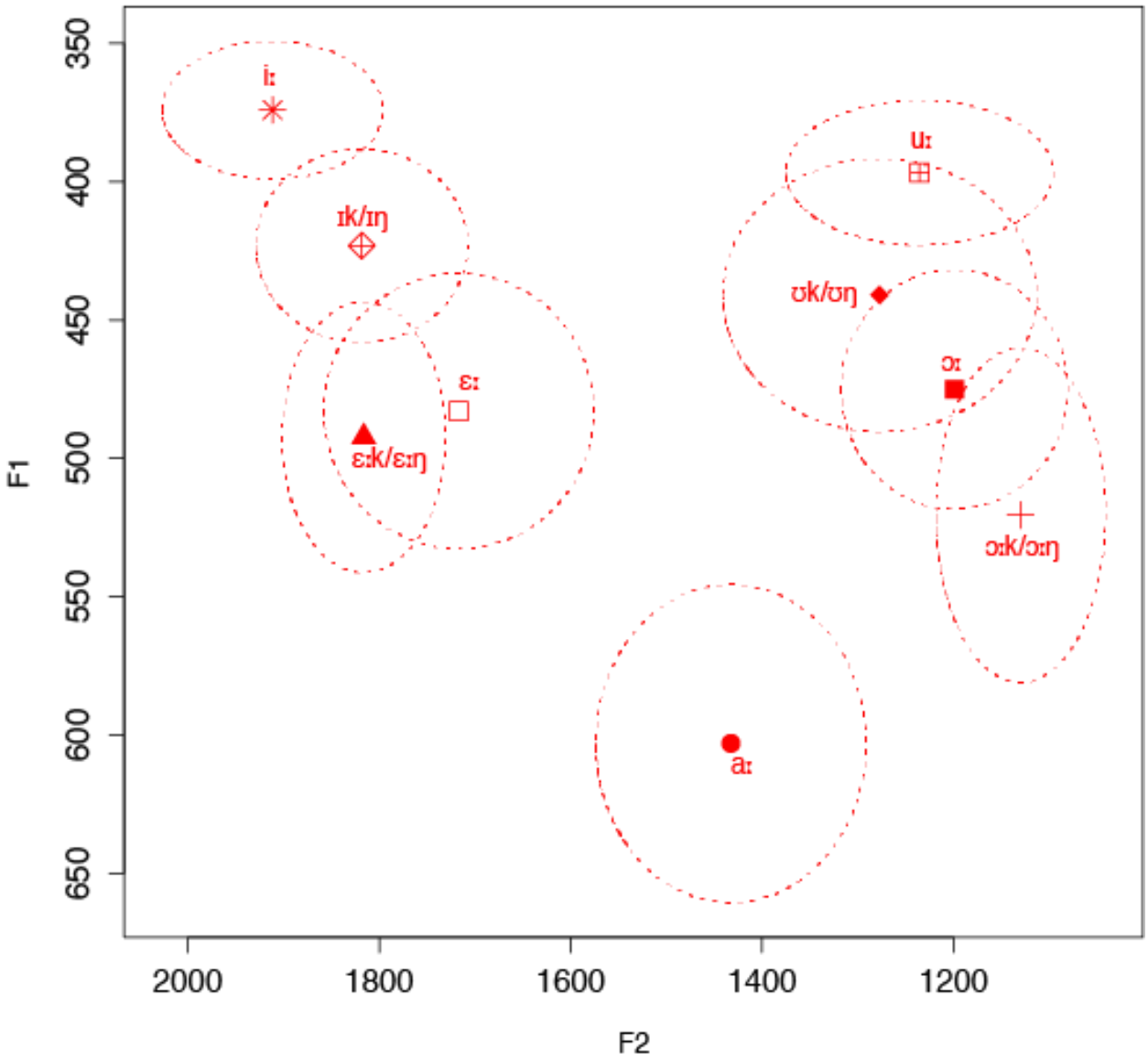

Figure 6. GEN 2 male speakers, Mean F1/F2 with ± 1 standard deviation, Normalized using the Lobanov method.

Abstract (Traditional Chinese)

本文通過分析語音材料，探索廣東話元音的社會語音學蘊涵。我們實驗採用的語音材料，來自多倫多繼承語的變體與變化語料庫（Heritage Language Variation and Change in Toronto Corpus）。本文重點分析在兩代發音人中，兩個對立元音（/iː/和/uː/）中的四個音位變體（[iː]、[ɪk/ɪŋ]、[uː]和[ʊk/ʊŋ]）的不同發音狀況。我們一共有 20 位發音人，每一位發音人都提供了 30 個元音發音。本文通過分析這 30 個元音發音的第一和第二共振峰頻率（F1 和 F2），探究以上所說的四個音位變體在不同發音人中的實際體現。分析結果表明：1）對於不同的世代（在不同的兩代人的發音中），音位/iː/和/uː/都有不同的變體，而這些不同的變體的分化條件在兩個世代中（兩代人中）都得到了保存；2）音位變體的實現與世代以及性別都有關聯。比如說，我們發現在第二代的女性發音人中，[ɪk/ɪŋ]的發音舌位最靠後，而[iː]的發音舌位最靠前。我們討論了產生這種結果的三個可能的原因，它們分別是：內部動因、語音同化、音系影響。繼承語音系學研究在當今是一個較為新興的方向。通過本文的分析與討論，我們展示了多層次的比較（包括跨世代、跨語言以及跨社區的比較）在繼承語音系學研究中的重要作用。

Author's address

Holman Tse
Department of Linguistics
University of Pittsburgh
2816 Cathedral of Learning
Pittsburgh, PA 15260

hbt3@pitt.edu