Beyond the Monolingual Core and out into the Wild: A Variationist Study of Early Bilingualism and Sound Change in Toronto Heritage Cantonese

by

Holman Tse

B. A. in Anthropology, University of California – Santa Cruz, 2003

M. A. in Linguistics, University of Chicago, 2005

Submitted to the Graduate Faculty of

The Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

PhD in Linguistics

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Holman Tse

It was defended on

September 14, 2018

and approved by

Shelome Gooden, PhD, Associate Professor, Linguistics

Alan Juffs, PhD, Professor, Linguistics

Naomi Nagy, PhD, Associate Professor, Linguistics, University of Toronto

Dissertation Advisor: Scott F. Kiesling, PhD, Professor, Linguistics

Beyond the Monolingual Core and out into the Wild: A Variationist Study of Early Bilingualism and Sound Change in Toronto Heritage Cantonese

Holman Tse, PhD

University of Pittsburgh, 2018

This dissertation focuses on variation and change in the vowel system of Toronto Heritage Cantonese with the goal of pushing variationist research on sound change beyond its monolingually oriented core (Nagy 2016) and in approaching the study of heritage languages from the perspective of spontaneous speech. It addresses the possibility of contact-induced inter-generational vowel shifts, mergers, and splits in native vocabulary. It also addresses the extent to which demographic, ethnic orientation, or language use factors may account for these changes. The data comes from the Heritage Language Variation and Change in Toronto Project (Nagy 2011) and includes hour-long sociolinguistic interviews from Toronto residents of different age, sex, and generational backgrounds speaking in Cantonese along with Ethnic Orientation Questionnaire data and a picture description task from each speaker. The mean F1/F2 of each vowel category from each of 32 speakers were measured in native (and integrated English) vocabulary. The results show lack of vowel shifts, evidence for merger in progress of /y/ ~ /u/, and evidence for a pre-nasal split in /ɛ/. The speakers who lead in this merger and split are the ones who used the least amount of Cantonese in the interview samples. The lack of the same structural changes from Hong Kong speakers further supports an account based on contact-induced change. These findings challenge Labov's (2007) Transmission and Diffusion model and suggest more sociolinguistic engagement with theoretical models of contact-induced change (cf. Thomason & Kaufman 1988, van Coetsem 2000).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

As is the case for just about anyone working on a doctorate degree, this has been a long journey for me. Needless to say, such a long journey would not have been possible without the help and support of so many people I have encountered along the way.

First and foremost, I must thank each and every individual member of my committee. To Scott, I thank you for all of the support and guidance you have provided me over the years from the moment I was admitted to the PhD program to the completion and submission of this dissertation. I am grateful for your breadth of knowledge of everything that falls under the umbrella of "sociolinguistics". To Shelome, I thank you for your classes on language contact and sociophonetics, for our many spontaneous conversations, and for making me realize that maybe I am on the path to becoming an early 21st Century Weinreich. To Alan, I thank you for your Research Methods course, your expertise in linguistic theory and acquisition, and your expertise on a "related language". 我真的太谢谢你! And to Naomi, I thank you for inspiring me to work on the HLVC Project, for arranging my visit to Toronto in 2014, for your leadership in developing a very innovative approach to language contact, and for your continued support of my academic and professional development.

Next, I must acknowledge the sources of funding supporting my project. 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. My dissertation research was funded by an Andrew Mellon Pre-doctoral Fellowship (2015-16), a Teaching Fellowship (2016-17), a Summer Research Fellowship (2017), and the Dean's Tuition Scholarship (2017, 2018) all from the University of Pittsburgh Dietrich School of Arts & Sciences.

The HLVC Project would not have been possible without the help of many research assistants and of course the interviewees themselves. Although I have not met everyone, I am indebted to all of those involved in the project for making my dissertation possible. I would especially like to thank: Naomi Cui, Minyi Zhu, and Vina Law for getting the computational logistics of my project started; Andrew Peters for continuing their work; Sam Lo for getting Homeland data together for me at the right time; Abigail Chan, Rachel Coulter, and Natalie Cheuk for their contributions as University of Pittsburgh undergraduate research assistants.

Networking and presenting my research at various stages has also been helpful in the development of my dissertation. I would like to thank everyone in the University of Toronto Linguistics Department for welcoming me during my short stay in Toronto and for our many conversations about linguistics in Canada. Earlier versions of my work were presented at CVC 8 in Kingston, ON, NWAV 44 in Toronto, APLA 39 ALPA in St. John's, NL, WICL 3 in Columbus, OH, LSA 2016 in Washington, DC, LSA 2017 in Austin, TX, and the Workshop on Variation and Change in Chinese in Melbourne via Skype. I would like to thank audience members at these conferences for their interest and feedback.

In my home department, there are also many other individuals I must thank. For providing feedback on my project at various stages, whether with one-on-one conversations or in response to various presentations, I would like to thank Karen Park, Matthew Kanwit, Jevon

Heath, Melinda Fricke, and Abdesalam Soudi. The resources and computational expertise available in our Language Media Center were also helpful at various stages of my project. For that I must thank Na-Rae Han and LaShanda Lemmon. Although not directly involved with this project, David Mortensen also deserves special acknowledgement for supporting me and my earlier work at Pitt. Many ideas presented in this dissertation are a continuation of conversations I have had with him over the years. I would also like to thank Marta Ortega-Llebaria for deepening my knowledge of acoustic phonetics. That, of course, is a very central part of this dissertation. Last, but not least are my fellow graduate students. Whether they are still at Pitt or have already moved on, they have kept me company over the years. I would especially like to thank Matthew Eppley, Matthew Hadodo, Miroo Lee, Cass Lowry, Farrah Neumann, Sean Nonnenmacher, Joe Patrick, Héctor Ramírez-Cruz, Alexandra Szczupak, Anthony Verardi, and Zhaohong Wu for being there during some of the most challenging moments of dissertation work.

Much of the write up of the results and analyses was done in San Francisco. There, I must thank the local chapter of the Shut Up and Write group for keeping me focused and both Dominic Yu and Patrick Chew for their amazing expertise in Cantonese linguistics. At times I was able to write in Pittsburgh, I would like to thank Ellen Smith for organizing the Dissertation Writers Boot Camp and the many Writer's Days in the Writing Center.

Many of the ideas discussed in this dissertation did not start with my work on the HLVC Project. They have a much longer history that stretches back to the beginnings of my post-secondary academic career. Judith Aissen receives credit for being my first ever linguistics instructor at UCSC and for making me realize that I really do know a lot about a language that I did not learn how to read or write. At UCSC, I was fortunate to have been taught by an amazing group of faculty and graduate student TA’s. Jaye Padgett, Adam Albright, Armin Mester, Andy Wedel, Nathan Sanders, and Vera Lee-Schoenfeld all deserve credit for getting me hooked on phonetics and phonology. But of course, at UCSC, anthropology was my actual major. In that department, I was also fortunate to learn from an amazing group of faculty. I must especially thank Don Brenneis for being a mentor and for introducing me to linguistic anthropology. Upon reflection, I realize that my current interest in language contact really traces itself back to my interest in culture contact, a topic that deeply interested me in my cultural anthropology courses. My views of the social world remains rooted in anthropological perspectives.

At the University of Chicago, I would like to thank Alan Yu for introducing me to acoustic phonetics and to Praat, which of course is one of the most important tools I used in this dissertation study. I would also like to thank Michael Silverstein for our many deeply intellectual conversations and for challenging me to think more deeply about language. The LSA Institute at UC Berkeley in 2009 was an important moment for me. There, I found the encouragement and inspiration to continue the academic path I started after taking a four year detour into the working world. I would especially like to thank Lauren Hall-Lew and Paul Kerswill whose work on San Francisco and London respectively gave me the motivation to come to Pittsburgh for a PhD. I do not regret that decision. Since then, the city of Pittsburgh and its rich sociolinguistic history and unique dialect have supported the continued development of my interest in language and culture contact.

Finally, I must thank my entire family including parents, sister, aunts, uncles, cousins, and especially my late grandmother not only for all of their love and support over the years but also for making me a heritage Cantonese speaker. This dissertation is about people who are linguistically like myself.

1.0 INTRODUCTION

At a time when generative synchronic linguistics was sweeping the world of linguistics, Labov famously said, “I have resisted the term sociolinguistics for many years, since it implies that there can be a successful linguistic theory or practice which is not social” (1972:xiii). More than 40 years later, sociolinguistics has become firmly implanted as a subfield of linguistics widely recognized as distinct from other subfields. The increasing specialization of research has even made it possible to talk about (quantitative) variationist sociolinguistics (Tagliamonte 2011) as opposed to interactional sociolinguistics as well as to talk about inter-subfield endeavors such as sociophonetics (Thomas 2011). “Three waves” of variationist research have also since emerged (Eckert 2012), each contributing to the continuing development of a subfield that was not originally intended to be a distinct subfield. Though some have argued for even earlier roots in both the US (Shuy 2003) and in Europe (Calvet 2003), my point here is to illustrate the simultaneous breadth and specialization of what researchers now consider sociolinguistics.

In this dissertation, I contribute to what is now recognized as (quantitative) variationist sociolinguistics while also echoing Labov’s initial resistance against “sociolinguistics” by engaging with theoretical frameworks and ideas that are now widely seen as outside the subfield but that are crucial to understanding variation and change in the particular community I examine. The specific topic is vowel variation and change across two generations of Cantonese speakers in

Toronto, Canada. These two generations include an immigrant generation (GEN 1) and a Canadian raised generation (GEN 2). The transition between these two generations involves a transition from being dominant in Cantonese (with variable knowledge of English and Mandarin) to a generation dominant in English. To strengthen or refute an account based on contact-induced change, I also analyze a set of speakers from the Homeland (Hong Kong) to determine whether or not changes innovated by Toronto speakers are unique to Toronto.

This topic involves multilingualism, though the focus will largely be on only two of these languages, Cantonese and English, with an even more specific focus on the possibility of English to Cantonese influence rather than Cantonese to English influence. The theoretical topic is sound change. The most well developed framework for sound change in the variationist sociolinguistics paradigm, the Transmission & Diffusion Model, is not equipped to handle sound change in such a setting. For this reason, I draw heavily from and engage with experimental research on heritage language phonetics and phonology and models from contact linguistics. Linguistics research on heritage languages has been described as a relatively new field (Polinsky and Kagan 2007). Thus, the way I resist "sociolinguistics" is by bringing together one of the newest research areas (the study of heritage languages) with one of the oldest topics in modern linguistics (the study of sound change).

The goals of this dissertation are summarized in the title. They are as follows:

1) To push "variationist research beyond its monolingually-oriented core" (Nagy 2016:24).

2) To push research on heritage language (to be defined below) phonetics and phonology out into "the wild"[1] .
3) To motivate a research program (based on the Uniformitarian Principle) that addresses the implications of early bilingualism (defined below) to sound change theory.

The first goal is one shared by Nagy, who describes the goal of "pushing variationist research beyond its monolingually-oriented core" (2016:24). As Nagy and Meyerhoff (2008) have observed, variationist research has a strong monolingual bias. One of the biggest research gaps is the lack of variety of languages and types of communities investigated. This stands in contrast to many other subfields in which interest in linguistic diversity has grown in the past two decades. This research gap has inspired the development of the Heritage Language Variation and Change in Toronto (HLVC) Project (Nagy et al. 2009; Nagy 2011; Nagy 2016). The data for this dissertation comes from the HLVC Project Corpus.

The second goal is to bring sociolinguistics research methodology to the study of heritage language phonetics and phonology. The study of heritage languages is a relatively new topic of interest among linguists. Most research on heritage languages, however, has focused on heritage speakers in classroom settings or in other controlled conditions rather than in the naturalistic contexts ("in the wild") favored by sociolinguistics research. In this sense, the "social" aspects of language are still very central to this dissertation. At the same time, research on heritage speakers need to be seriously considered even if many of these studies are based on controlled settings. Lynch and Polinsky (2018) have discussed how heritage speakers challenge traditional ideas of

---

[1] This metaphor is inspired by the title of Polinsky and Kagan's (2007) article *Heritage Languages: In the 'Wild' and in the Classroom*.

native speaker hood that are still pervasive in linguistic theory. Labov's Transmission and Diffusion Model (2007), for example, is based on a strict distinction between the linguistic behavior of native and non-native speakers of a language.

The third goal is a long-term goal that involves extending Labov's use of the Uniformitarian Principle to the study of vowel variation and change in multilingual communities. In presenting the Uniformitarian Principle, Labov, Yaeger, and Steiner (1972) say "that the forces which produce sound change today are the same as those which operated to produce the historical record" (1972:1). Thus, how can we better understand sound change in bilingual and multilingual communities of the past by better understanding sound change in bilingual and multilingual communities in the present? While historic multilingual settings have certainly been addressed in creole linguistics, relatively little variationist research has investigated the actuation and propagation of sound change (particularly for vowels) in progress in all possible directions (ex: from dominant language to heritage language vs from heritage language to dominant language). Heritage language communities provide examples of directly observable multilingualism in the present. Yet, few have considered studying such communities with the goal of better understanding sound change. If Labov's original goal was to better understand the past by studying the present, variationists must be better equipped to understand different types of multilingual settings in the present. As Gooden notes "the fields of sociolinguistics and language contact have wafted in and out of love over the years since Weinreich (1953)" (in press). Although variationist research on communities characterized by language contact is not new, "scholars in the subfields have operated largely in different arenas" (Ibid.). Thus, I attempt to bridge the gap between these two subfields because "without looking at language contact, we might … lose sight of the long-term processes of language variation and change which give linguistic variables history" (Ibid).

Before elaborating further on the details of this dissertation, I discuss two important terms: *early bilingualism* and *heritage language*. I define *early bilingualism* as *a form of individual speaker bilingualism in which two languages are learned prior to adolescence*. What counts as having learned or acquired a language can be a contentious issue. For the purpose of this dissertation, I define having learned a language as knowing a language well enough to construct novel sentences in at least a few spontaneous speech contexts with communicative intent. Thus, an *early bilingual*, according to my definition, would still be an early bilingual even if that person's usage of the less proficient language is limited to contexts involving speaking with family members. Proficiency in each language may also be balanced, but that is not a requirement. All that matters in this definition is that an individual has learned two languages prior to adolescence.

Nagy defines *heritage languages* "in the Canadian context as mother tongues other than Canada's two official languages (English and French)" (2016:16). This definition generally coincides with the Canadian census definition, which makes a distinction between heritage languages spoken by immigrant groups and indigenous languages spoken by pre-colonial groups. Montrul defines heritage languages more broadly by describing them as "culturally or ethnolinguistically minority languages that develop in a bilingual setting where another sociopolitically majority language is spoken" (Montrul 2015:2). Under this definition, indigenous languages would be considered heritage languages unlike under the Canadian census definition.

For the purpose of this dissertation, either Nagy's (2016) or Montrul's (2015) definition would be suitable. The advantage of Montrul's definition is that it is broader making it possible to discuss Toronto Heritage Cantonese in comparison to similar sociolinguistic settings elsewhere in the world. At the same time, the findings from this dissertation may be limited to the specific sociolinguistic setting under discussion. For this reason Nagy's (2016) Canadian-specific definition would also be a suitable definition. This dissertation is a starting point for building more generalizable claims about sound change in the context of early bilingualism and similar sociolinguistic settings.

More prototypical examples of heritage languages, as discussed in the growing literature on this topic, include cases involving languages spoken in a diasporic context. In other words, heritage languages are typically tied to a social history of migration from a "homeland" to a "host society". Of both descriptive and theoretical interest for researchers is the linguistic outcome that results from the transition from the first generation (the immigrant generation born and raised in the homeland) to the second generation (born and raised in the host society). The two generational groups, which I will refer to as GEN 1 (first generation) and GEN 2 (second generation), are differentiated by relative age of acquisition of the host society's dominant language although the first generation does not always master this language. The transition from GEN 1 to GEN 2 is, thus, characterized as a change from monolingualism (or adult L2 acquisition) to early bilingualism. A social change resulting in change in the relative age of acquisition of different languages brings psycholinguistic considerations together with sociolinguistic ones. This dissertation topic is, thus, not one that can be considered within the exclusive domain of sociolinguistics.

The specific linguistic features analyzed in this dissertation are consistent with the methodology and analytical procedures pioneered by Labov, Yaeger, and Steiner (1972). The focus is on the vowel system and whether or not a sociolinguistic change involving an entire generation learning a new language at an early age can lead to change in the first language spoken within such a community. The research questions are framed in terms of the typology of vowel changes described in variationist research. In other words, are there chain shifts, mergers, and/or splits?

The specific research questions are stated as follows:

(Q1) Is there evidence for contact-induced inter-generational vowel shifting in native vocabulary?

(Q2) Is there evidence for contact-induced vowel mergers or vowel splits in native vocabulary? Four specific parts of the Cantonese vowel space are addressed:

(Q2a) Is there evidence for a merger between /y/ and /u/?

(Q2b) Is there evidence for an increasing acoustic split between /i/ and /ɪ/?

(Q2c) Is there evidence for an allophonic split in /ɛ/?

(Q2d) Is there evidence for an allophonic split in /ɔ/?

(Q3) To what extent can demographic (sex, age), ethnic orientation (Ethnic Orientation Questionnaire or EOQ scores, individual EOQ responses), Cantonese % Scores, Cantonese Word Count Scores, and English Word Count Scores account for the propagation of the specific shifts, mergers, and splits observed in the data?

By addressing these descriptive questions about the data, the theoretical question addressed in this dissertation is as follows:

(T1) What are the implications of the findings from this study for models of contact-induced sound change?

This dissertation contains eight chapters. I describe the contents of these chapters in the paragraphs below.

In Chapter 2, I discuss the research literature on sound change from its beginnings in historical linguistics to more recent work that has developed within the variationist sociolinguistics paradigm. Of central importance in this chapter is Labov's (2007) Transmission and Diffusion Model and how it developed based on more than 40 years of research on sound change in progress. This is the most sophisticated model of sound change developed within the variationist paradigm. Yet, as I will argue, it is also based on several assumptions that are problematic for the study of phonetic variation and change in communities characterized by early bilingualism. Some of these problematic assumptions have also been critiqued by Third Wave researchers (Eckert 2012). I review Third Wave studies that have focused on multilingual contact settings and discuss how heritage language contact settings support Third Wave critiques of the speech community and the vernacular. This dissertation contributes to this ongoing discussion by focusing on the potential influence of a dominant language on a heritage language when both languages are acquired during childhood.

In Chapter 3, I present two models of contact-induced change: the Thomason and Kaufman (1988) Model (henceforth "TK") and the van Coetsem (1988; 2000) Model (henceforth "VC"). I discuss three major problems that the current dissertation will attempt to resolve. The first is skepticism against claims about contact-induced change. I discuss how methodological problems may have led Labov (2008) and many historical linguists to be skeptical of its importance. If the outcome of internally motivated change can be identical to the outcome of externally motivated change involving structural influence (as predicted by both the TK and VC Models), how can the two processes be distinguished from each other?

Although both the TK and VC Models appear to be more suitable models for the study of heritage language change, it is unclear if either model is a better model because of the different, but not completely mutually exclusive, perspectives taken by each model. While the former takes the perspective of socio-historical continuity, the latter takes the perspective of the cognitive processes operating in individual speakers. Another important difference is the issue of relative age of acquisition vs. linguistic dominance. In Chapter 2, heritage language speakers are described as problematic for the Transmission and Diffusion (henceforth "TD") Model (Labov 2007) because they typically become dominant in their second language. For this reason, the van Coetsem Model appears to be a more appropriate model because it addresses this issue by focusing on individual speaker linguistic dominance as the underlying basis between different contact mechanisms. It predicts that a speaker's dominant language can have structural influence on a speaker's non-dominant language and that this influence decreases as proficiency in both languages becomes more balanced. Many studies of heritage language vowels based on experimental methods, however, show stability in terms of phonemic contrasts and that this may be due to early acquisition of two languages even if speakers subsequently become dominant in a different language (see for example Yang 2014; Mack 1990; Godson 2004; Saadah 2011; Baker and Trofimovich 2005; Chang et al. 2011). These same studies, however, also show evidence of dominant language influence on the phonetic production patterns of heritage speakers. This creates a puzzle as to how phonetics and phonology may be related to each other under such contact settings. Could these low-level phonetic differences be merely phonetic or are they also precursors to sound change?

Although the VC Model does seem more promising in its recognition of how linguistic dominance does not necessarily correspond with one's first language, the TK Model does not appear to be completely wrong, but rather incomplete. The TK Model does not predict that all individuals within a heritage language community will be equally influenced by the dominant language. Rather it predicts that some members will be and that over time, the features of the innovative speakers will spread to the rest of the community. Furthermore the TK Model includes a caveat that given the right social context, language ideologies and attitudes can override any of the tendencies predicted by this model. This seems consistent with the Third Wave (Eckert 2012) focus on local social meaning. What is not clear is the relationship between individual speaker actuation of change and its propagation across a community. I discuss how a variationist approach following Nagy (2011; 2015) can address problems and unanswered questions that have developed from the models and experimental studies presented in this chapter. These problems include 1) the methodological issues of proving contact-induced change, 2) the relationship between the individual and the community in sound change, and 3) the interpretation of low-level phonetic differences among early bilingual speakers. Following (Nagy's 2011) protocol for strengthening accounts based on contact-induced change, this study includes 1) an inter-generational comparison of Toronto speakers (GEN 1 vs. GEN 2), 2) a diatopic comparison (Toronto vs. Hong Kong), and 3) a cross-linguistic comparison (Toronto Cantonese vs. Toronto English). Thus, while the central focus is on the inter-generational comparison, data from Hong Kong are also included to strengthen or refute an account based on contact-induced change.

In Chapter 4, I present the socio-historical context of Cantonese both in Hong Kong and in Toronto. I discuss how the economic growth of Hong Kong is tied to the development of

Hong Kong Cantonese as the prestige variety of Cantonese. Toronto's Cantonese community grew beginning in the 1960s when Canadian immigration laws were loosened. The changing political climate in Hong Kong in the 1980s and 1990s led to major waves of migration to Canada, with many immigrants ending up in Toronto. Cantonese is now the second most spoken mother tongue in the Greater Toronto Area. Although only 5% of the population in the Greater Toronto Area speaks Cantonese as a mother tongue, language maintenance has been supported by the economic development of many Chinese-owned businesses and by Canada's multiculturalism policy.

In Chapter 5, I discuss the specific linguistic issues that are the focus of this dissertation as well as the data and methodology. This includes discussion of Cantonese vowel phonology. Many different analyses of the Cantonese vowel system have proposed in the literature. For this study, I follow Zee's (1999) analysis which describes a system with 11 monophthongs and 11 diphthongs. The focus of this dissertation is on the monophthongs. The monophthongs can be further subdivided into a tense system and a lax system. The tense vowels include: /i/, /y/, /u/, /œ/, /ɛ/, /ɔ/, and /a/. The lax vowels include /ɪ/, /ʊ/, /ɵ/, and /ɐ/. After describing the distribution of these vowels and the tone categories of Cantonese, I discuss specific hypotheses about possible vowel shifts, mergers, and splits that might develop. The data for this dissertation comes from the Heritage Language Variation and Change (HLVC) in Toronto Project. The corpus from this project includes transcribed hour-long sociolinguistic interviews from Toronto residents of different age, sex, and generational backgrounds speaking various heritage languages along with Ethnic Orientation Questionnaire data and a word list task from each speaker. Cantonese is one of eight languages that are part of the growing corpus. I discuss what the responses to the Ethnic Orientation Questionnaire show about the speakers being analyzed and how these responses can be used to model variation in the data. I also discuss a set of language uses scores that I have created to address the extent to which individual speaker proficiency in Cantonese might also account vowel variation and change. The rest of this chapter is devoted to discussion of the methodological procedures I used to address the research questions.

In Chapter 6, I present the results of this dissertation, which are based on a total of 33,179 vowel tokens. In terms of vowel shifting (Q1), the GEN 2 group is more conservative than the Homeland group. While the Homeland results show evidence for vowel shifting for four different vowel categories (/i/, /ɪ/, /ʊ/, and /ɔ/), only one vowel showed evidence of an inter-generational shift in Toronto: /y/. In terms of mergers and splits (Q2), however, the GEN 2 group appears to be more innovative. A closer analysis of /y/ shows that its retraction is part of a change towards merger of /y/ and /u/, which would be a change expected due to structural influence from Toronto English, which lacks a contrast between two high round tense vowels. The lack of the same change in Hong Kong provides further support that it is a contact-induced change. The factors that account for the /y/ retraction are also factors consistent with a contact-induced change account. The other major structural change identified in the GEN 2 group is the two-way allophonic splitting of /ɛ/. This vowel is fronted preceding nasals, which could be due to phonetic similarity with the pre-nasal tense variant of Toronto English /æ/, which is raised as it is across many dialects of North American English. Before stops, Cantonese /ɛ/ is retracted and thus appears to converge with the pronunciation of Toronto English /ɛ/, which is retracted and part of the Canadian Vowel Shift. Too little data, however, is available on the retraction of Cantonese /ɛ/ to determine what factors facilitate this change. For pre-nasal /ɛ/ fronting, the only significant predictors were CAN % Score and ENG WC Score. These factors and the absence of the same change in Hong Kong support a contact-induced change account. The other two parts of the vowel system analyzed did not show evidence of any GEN 2 change.

In Chapter 7, I account for the findings reported in Chapter 6. I focus the discussion on two broad patterns. The first is the lack of vowel shifts while the second is the fact that the few changes that have been identified are structurally-motivated changes. I discuss several factors that facilitate the general lack of change. Early acquisition of two genetically and typologically distinct languages is only part of the picture. Also important to consider are language use practices that have developed among Toronto Cantonese speakers. I discuss how GEN 2 speakers view Cantonese as a language that is inherently full of English loan words that are pronounced with Cantonese phonology. This awareness makes possible the use of correspondence rules (Thomason 2007), which refer to the use of sound correspondences between two different languages to create novel forms in one of these languages. Although this process leads to lexical innovation, I argue that it also reinforces phonological maintenance.

Although there are social factors that support phonological maintenance, structural changes still occur for some speakers. On the group level, the specific changes that do occur can be accounted for in terms of functional load. Vowel contrasts that have a small functional load are more susceptible to change than those that have a higher functional load. This explains how a change towards merger of /y/ and /u/ can develop. Vowel splits influenced by two phonetically similar allophones in the dominant language can also develop for vowel contexts that do not occur very frequently. On the inter-speaker level, factors that relate to language use are the best predictors that account for which individual speakers are the most likely to innovate these changes. I also discuss the theoretical implications of these findings. The TK and VC Models are clearly better equipped to handle the observed findings than is the case for the TD Model, but they also take different perspectives.

In Chapter 8, I conclude by discussing the broader implications of this study. This study pushes variationist research beyond its monolingually oriented core (Nagy 2016) by focusing on sound change in a heritage language. This study also offers a variationist perspective to the study of heritage languages, a research area that has been dominated by experimental approaches. Heritage languages are in a unique position to address many broader questions relevant to sociolinguistic theory due to the social context in which they develop and in which they are spoken. Guiding this project is the Uniformitarian Principle. If we recognize that multilingualism has historically and still is the norm throughout much of human civilization, then we need to develop a theory of sound change that acknowledges a broader range of possibilities created by diverse contact settings. A skeptic might ask why one would even bother studying sound change in a heritage language if people stop speaking the heritage language. On the topic of language loss in historical linguistics, Simpson say, "a common end-state of language attrition is disappearance of the old language as people shift to speaking another language. But the shift could be halted" (Simpson 2014:551). Simpson continues by saying that this is what may have happened to English immediately after the Norman Conquest. This, of course, is a controversial issue but such controversies about historic contact settings only further necessitate studies of similar contact settings in the present. The extent to which Toronto Heritage Cantonese may be similar to post-Norman Conquest English is debatable, but what is certain is that it is a directly observable language for researchers in the early 21st century. I hope this dissertation study encourages more research on sound change in different contact settings in the present so that we can better understand how sound change may have developed under different types of contact settings in the past.

2.0 THE GENEALOGY OF SOUND CHANGE RESEARCH

In this chapter, I trace the history of ways of thinking about sound change and contact-induced change and how these different ways of thinking have led to the current state of affairs of sound change research in variationist sociolinguistics. I begin in Section 2.1 by discussing the Neogrammarian Controversy, which was a debate between two models of change: The Family Tree Model and the Wave Model. From the 1960s onward, a variety of new ideas and approaches entered the scene. In Section 2.2, I discuss innovations in laboratory phonetics, which have made it possible to observe sound change in progress, and the Lexical Diffusion model (Wang 1969), which Labov described as a "recasted" (1994:10) form of the Wave Model. In Section 2.3, I discuss the framework presented in Weinreich, Labov, and Herzog (1968) and how that eventually paved the way to Labov's reconciliatory reformulation of the Neogrammarian Controversy as a distinction between *transmission* and *diffusion*. In Section 2.4, I discuss Third Wave critiques of two concepts central to the Labovian framework: the speech community and the vernacular. I will explain how these two concepts are problematic for the study of heritage languages (as an L1) spoken in a context involving an inter-generational transition to a different dominant language (as an L2). I summarize the main points of this chapter in Section 2.5.

2.1 THE NEOGRAMMARIAN CONTROVERSY

The concept of sound change was first proposed by the Neogrammarians of the 19th century. By comparing the pronunciation of vocabulary across different ancient languages, philologists discovered systematic sound correspondences as inferred through written evidence. Based on these correspondences, it became possible to reconstruct Proto-Indo-European, the ancestor language of many modern European languages. It is here that we see one of the first major breakthroughs in the history of modern linguistics. If it is possible to reconstruct the genetic lineage of modern-day languages through the identification of sound correspondences, then there must have been a set of "laws" that led to the development of such patterns. This became known as sound change. The idea that sound change is a regular and exceptionless process became part of the Neogrammarian Hypothesis. The Family Tree Model for the genetic classification of languages became closely tied to this hypothesis since it was an assumption that made the reconstruction of language family trees possible via the Comparative Method.

The Neogrammarian view, however, was not universally accepted. Nineteenth century dialectologists identified many examples that they claimed challenged the supposed regularity of sound change. So began the Neogrammarian Controversy, which Thomas describes as "the first great controversy in linguistics" (2011:4). The alternative explanation proposed was the Wave Model. According to the Wave Model, changes start with individual words in one geographical location and then they spread to other locations just as a wave does. Different words may spread in different directions leading to different patterns of variation in the dialects spoken in different places. The outcome is summarized in the widely cited slogan, "every word has its own history" (Campbell 1999).

One example discussed by Campbell (1999:189–190) to illustrate the Wave Model is French dialects spoken in Normandy in the 19th century. While Latin /k/ became /ʃ/ in Standard French, this change did not affect all cases of inherited /k/ in all Norman dialects. According to the Wave Model explanation, whether a word maintained the /k/ or adopted the /ʃ/ pronunciation depended on whether the meaning of the word was more closely linked to rural life or to church life. For example, words that have an association with the church such as [kandel] > [ʃandel] "candle" and [kante] > [ʃante] "to sing" were argued to have been pronunciations diffused by priests who adopted the Parisian pronunciations of these words. Because of their particular history, these words were affected by the change of /k/ > /ʃ/, which started in Paris and spread across the region from there. Other words in these Norman dialects, however, retained the /k/ pronunciation because of their association with rural life. Some examples of this include [kattu] 'cat' (instead of [ʃat] as in Standard French) and [kampus] 'field' (instead of [ʃam] as in Standard French). Some small pockets in the region that were less influenced by Parisian priests retained /k/ for a larger set of words. Whether words retained Latin /k/ or had /ʃ/, thus, depended on the particular history of the particular words involved. Each word has its own history and its history depends on the socio-historical contexts of its usage. As opposed to the Neogrammarian Model of sound change, the Wave Model assumed that change is socially or externally motivated.

In response to Wave Model proponents, the Neogrammarians offered other explanations to account for these "exceptions." Bloomfield (1933) referred to these exceptions as "residual forms," thus highlighting the Neogrammarian view that they are not exceptions to sound change but rather remnants from two irregular processes. One is borrowing while the other is analogical change. Both result in innovations.

Borrowing can occur with vocabulary from a genetically distinct language. In many of these cases, it is clear that these words are borrowed because they introduce phonotactic patterns that are not found in the borrowing language. Cases of dialect borrowing, however, can be more difficult to recognize on the surface. This would apply to the Norman French /k/ ~ /ʃ/ problem. If words are borrowed from another dialect of the same language, these words can give the impression that they are exceptions to general patterns when in fact they are simply borrowed from a dialect that has developed different patterns. This is because different dialects of a language include many words with similar forms that may have undergone different sound changes in their historical development even though they look similar enough in form to be considered cognates. Thus, in recognizing some forms as borrowed from a related dialect, the Neogrammarians strengthened their view of exceptionlessness in sound change by identifying loan words as not subject to the same process. Following Neogrammarian reasoning, changes that do not fit existing patterns in a language must be due to some sort of external force such as contact with speakers from a different community.

The other path that leads to innovation according to the Neogrammarians is analogical change. An example of analogical change is in the development of the English plural for "cows" from Old English ([kyː] > [kawz]). This appears to be an exception to the general correspondence between Old English [yː] with Modern English [aj] as illustrated in Table 1. The singular form "cow", however, does follow a general correspondence between Old English [uː] and Modern English [aw] as shown in Table 2. According to Bloomfield, the form "cows" appears beginning in 1607 when an intermediate form, "kine" (presumably [kajn]), was also in use (1933:404). By this time, English had developed the use of the plural /-s/ suffix as the most common way of forming plurals. This contrasted with earlier stages of the language when the regular plural formation process involved vowel alternations. This earlier process is retained in Modern English in the plural of "mouse" ("mice"), which is also included in Table 1 and Table 2. Thus, by 1607, some speakers developed the more morphologically regular form [kawz] to replace the morphologically irregular form [kajn]. This innovation involves adding the plural morpheme /-s/ and phonetically conditioned voicing to the pronunciation of the singular form. The development of this regularized form is based on analogy with other examples of plural forms in English that involve the same process.

Table 1. Old English [y:] ~ Modern English [aj] Correspondences

Old English Forms	Modern English Forms
[hwy:]	'why' [(h)waj]
[my:s]	'mice' [majs]
[bry:d]	'bride' [brajd]
[ky:]	**'cows' [kawz]**

Table 2. Old English [u:] ~ Modern English [aw] Correspondences

Old English Forms	Modern English Forms
[hu:]	'house' [haws]
[mu:s]	'mouse' [maws]
[u:t]	'out' [awt]
[ku:]	**'cow' [kaw]**

By comparing and contrasting the development of the singular and plural forms for both "mouse" and "cow", we can see an illustration of Sturtevant's Paradox. According to Sturtevant, "phonetic laws are regular but produce irregularities. Analogic creation is irregular but produces regularity" (Sturtevant 1947:109). In Table 1 and Table 2, we can see that phonetic change is more regular (in the Neogrammarian sense) in its across-the-board effects. Yet, the results are irregular in terms of irregular allomorphs such as in “mouse”/“mice”. Analogic change, on the other hand, is irregular in terms of which words are affected. For example, there is no clear linguistic reason why the /-s/ plural was added to “cow”, but not to “mouse”. Over time, extension of the /-s/ plural suffix to a larger percentage of English vocabulary has created more overall regularity in the plural formation process. This example illustrates how analogic change “confirms the assumption of phonetic change” (Bloomfield 1933:405). We can see by way of contrast with analogical change that sound change is still described as an exceptionless process while the process of analogical change is less predictable in terms of which word it affects but does result in increased regularity in the morphology of a language.

One major problem that linguists of the 19th and early 20th centuries were unable to answer was the observational problem. The general consensus up through the 1950s was that sound change is too gradual of a process to be observed directly. Hockett says that while “borrowing and analogy can bring about marked reshaping of a single idiolect … Sound change … does not *noticeably* do this, though, in theory we must assume it does” (1958:444). He also stated explicitly that “the direct observation of sound change is impossible” (Hockett 1958:445). Bloomfield also noted that, “the occurrence of sound change, as defined by the neo-grammarians [sic], is not a fact of direct observation, but an assumption” (1933:364). In spite of the success of the Neogrammarian Hypothesis in accounting for many observed facts, historical linguists were unable to observe sound change actually happening. This left a big mystery that has since continued to drive research following many different approaches. The approach I adopt in this dissertation is the variationist sociolinguistics approach, which has shown that sound change can be observed. In the next section, I discuss two different approaches to addressing the observation problem that have influenced the variationist sociolinguistics approach.

2.2 ADDRESSING THE OBSERVATIONAL PROBLEM

From the 1960s onward, the Neogrammarian Controversy became reframed in response to new methodological innovations and theoretical ideas. In Section 2.2.1, I discuss the influence of laboratory-based phonetics research. I then discuss Wang's Lexical Diffusion Model in Section 2.2.2. I show in Section 2.3 how laboratory-based phonetics and the theorization of lexical diffusion as a mechanism for sound change would lead to the two parts of Labov's Transmission and Diffusion Model (2007; 2011). While the laboratory phonetics approach has provided further insight into the origins of Neogrammarian sound change (reframed as Transmission), the Lexical Diffusion Model (a reframed form of the Wave Model) has provided insight into a different process that could also account for some changes identifiable in the historical record.

2.2.1 The influence of laboratory phonetics on sound change research

Improvements in speech recording technology as well as other equipment to study the articulatory and physiological properties of speech and how speech signals are perceived have fostered the development of new methods to address the observational problem of Neogrammarian sound change. This body of research has focused on addressing underlying universal phonetic tendencies and constraints that could account for how sound change develops. If sound change really is the mechanical, phonetically continuous, and exceptionless process that the Neogrammarians assumed it to be, there must be some underlying mechanism driving it. This underlying mechanism must be something that is shared across all human beings. Following this line of logic, the place to look for the underlying mechanisms would be in the mechanics of speech production and perception. The common occurrence of the same processes across unrelated languages provided not only further support to the Neogrammarian Hypothesis but also provided a plausible explanation based on human physiological, articulatory, and perceptual constraints about why and how sound change could possibly originate.

Thomas (2011) discusses several examples of models that have developed based on experimental research. One issue that these models have raised is teleology. In other words, does sound change have a purpose? In Ohala's Listener Based Model, for example, sound change has no purpose. It is treated as an unintentional consequence that arises from misperception of co-articulated sounds. For Lindblom's H & H (Hypospeech and Hyperspeech) Model, however, there is a teleological origin to sound change based on variable communicative needs. Speakers vary between hypospeech (lack of attention paid to enunciation) and hyperspeech (enunciating more clearly) in everyday communicative interaction. Thus, innovations that arise from vacillation between different forms are the source of change according to the H & H Model. Blevin's CCC (Change, Chance, and Choice) Model incorporates Ohala's idea of misperception as "change", but also recognizes cases in which listeners interpret ambiguous phonetic signals. "Chance" refers to cases in which the listener picks the wrong phonological form when the phonetic signal is inherently ambiguous. Finally, "choice" is similar to Lindblom's idea of hypospeech vs. hyperspeech but involves cases in which listeners settle on a prototype sound that is different from that of the speaker.

In describing these three models, Thomas says, "they share a shortcoming. They all have a weak treatment of sociolinguistics" (2011:278). This could be because most of the research on sound change from a phonetics or phonological perspective has focused on "the origin of changes, as opposed to their propagation" (Thomas 2011:274). This assumes that "linguistic factors are responsible for the origin of changes and social factors are responsible for their spread" (*ibid.*). Thus, what has developed since the second half of the 20th century is a sub-field based division of labor with phoneticians and phonologists interested primarily in internally-motivated ('their origin') aspects of sound change while sociolinguists are primarily interested in externally-motivated ('their spread') aspects of sound change. Thus, there is not necessarily widespread disagreement among researchers about the existence of these two processes but rather a division of labor that has reinforced a theoretical distinction between these two processes. As I will explain shortly, Labov (2007; 2011), building on (Weinreich, Labov, and Herzog (1968), would later reinterpret this division as transmission vs. diffusion.

2.2.2 Wang's Lexical Diffusion Model

The Lexical Diffusion Model Wang (1969; 1979) presented an alternative explanation to the Neogrammarian problem of residues. It was an attempt to reconcile the Neogrammarian idea of exceptionlessness with the observation of what appear to be widespread exceptions. According to the lexical diffusion model, sound change (either phonetically abrupt or phonetically gradual) starts with individual words rather than sounds. Phonetic innovations in individual words then gradually diffuse across the lexicon to words that have the same sound in the same environment until the change reaches all words that have the original sound. Sometimes these changes are complete. These, Wang argued, would be the types of changes that the Neogrammarians described as regular and exceptionless. Other times, however, these changes are not brought to completion and are hence not completely diffused across the lexicon. Incomplete changes arise either because of competing processes that break the pattern or because of lack of time for a change to be completed. The state of a language at any given point in time, thus, could contain many residues of incomplete sound changes. Wang redefined a *regular sound change* as one in which "no other changes compete against it" (1969:10).

In explaining the rationale behind the model, Wang (1969) makes it clear that both phonetic and lexical issues need to be resolved in any theory of sound change. Both change in the overall phonetic pronunciation patterns and change in the specific words affected by pronunciation changes can be either gradual or abrupt. Wang (1969), thus, lists four possibilities. Sound changes on a person's vocabulary can be:

(1) phonetically abrupt and lexically abrupt,

(2) phonetically abrupt and lexically gradual,

(3) phonetically gradual and lexically abrupt, OR

(4) phonetically gradual and lexically gradual

An example of (1) would be if someone suddenly pronounces all instance of /p/ as /b/ across all lexical items that contain /p/. This is a possibility that Wang (1969) rules out because people do not suddenly and consistently change the pronunciation of all words in their vocabulary that contain a given sound[2]. (3) is essentially the Neogrammarian view. In other words, a phoneme changes gradually across all lexical items containing that phoneme. Lexical diffusion, however, involves either (2) or (4). For example, /ɔ/ becomes /ɑ/ (either as a phonetically gradual change or as an abrupt change from /ɔ/ to /ɑ/), but only in a subset of words (hence, lexically gradual) containing /ɔ/.

[2] Thomason's (2007) discussion of deliberate change and word games appears to present a counter-example.

Wang (1969:16) makes a distinction between lexical diffusion and dialect borrowing as two distinct processes but admits that there are often cases in which the two processes are difficult to distinguish. Dialect borrowing, he says, would involve borrowing the pronunciation of a word from another dialect while lexical diffusion would be a change in pronunciation that develops within the same dialect. Thus, to show lexical diffusion rather than dialect borrowing, one would have to show that the change is a change isolated to a specific dialect rather than a change that matches the pronunciation of a dialect in contact.

Although Labov has described this model as a “recasted” (Labov 1994:10) version of the Wave Model, Wang describes his model as a modification of the Neogrammarian Hypothesis (1969:10). Wang’s intent was a reconciliatory one. He saw value in the cross-linguistic typological findings from phonetics research as discussed in Section 2.2.1. Wang says that “The evidence is overwhelming that almost all sound changes come from an extremely small common inventory. Many of these changes are now beginning to find explanations in laboratory phonetics” (1969:22). Wang’s (1969) Lexical Diffusion Model was, thus, intended to complement rather than refute the growing body of research using experimental techniques. While the Neogrammarians argued that the phoneme was the basic unit of sound change and that sound change is phonetically gradual and lexically abrupt, Wang (1969; 1979) argued that the basic unit is the word and that sound change is lexically gradual. Lexical diffusion as a primary mechanism of change is a point that Labov would later critique.

2.3 FOUNDATIONS OF THE VARIATIONIST APPROACH

Variationist sociolinguistics developed partly as an alternative empirical approach to the observational problem of sound change in historical linguistics. In Section 2.3.1, I discuss the methodological innovations that have become the hallmark of the variationist approach to sound change. In Section 2.3.2, I discuss the concept of structured heterogeneity as presented in Weinreich, Labov, and Herzog (1968), a paper that has since become foundational to the subfield. I then discuss how Weinreich, Labov, and Herzog (1968) described multilingualism in terms of structured heterogeneity in Section 2.3.3.

2.3.1 Methodological innovations in the study of sound change in progress

Crucial to the empirical foundation of sociolinguistics is the use of audio recording technology, which was unavailable to the Neogrammarians of the 19th century. Recording technology has also seen constant development in the past several decades. This has been a benefit for phoneticians as well as for sociolinguists. Along with recording technology have come advances in statistical methods and in computational software, which have all made it possible to study sound change through the use of big data. Recording technology has also made it possible to analyze phonetic detail that may not necessarily be noticeable to the researcher (or to members of the speech community). The methodological approach made it possible to document "sound change in progress" (Labov, Yaeger, and Steiner 1972) by observing change in apparent time through the interviewing of speakers from different age cohorts in a given community. The Uniformitarian Principle, which states that the same processes that occur in the present are the same ones that applied in the past, was also important in rationalizing an approach to diachronic change based on present-day synchronic variation.

Also important to the variationist approach is the sociolinguistic interview. Unlike the research of phoneticians discussed in Section 2.2.1, the variationist approach favors naturalistic data obtained through spontaneous speech samples over speech obtained in controlled laboratory settings. The aim of a sociolinguistic interview “is to observe how people talk when they are not being observed” (Labov 1984:30). This has been referred to as the Observer’s Paradox (Labov 1984) since there is a clash between this goal and the fact that the presence of an interviewer inevitably creates a context in which formality is generally expected. To (at least partially) address the Observer’s Paradox, the interview asks questions designed to steer the interviewee towards use of more informal speech styles. This is achieved by identifying topics such as childhood, relationships, and community life. Interviewees are also allowed and even encouraged to go off tangent. It is not until towards the end of the interview that the interviewer asks questions about language or other topics that generally involve more attention paid to speech. Reading passages, minimal pair tests, and other linguistic tasks are conducted at the end of the interview. These linguistic tasks are important to include to address one of the disadvantages of focusing on informal spontaneous speech: lack of tokens of the specific linguistic variable being studied. For example, a specific study could be focused on a phonological context that does not occur often in spontaneous speech. Adding a minimal pair task at the end ensures that there are a minimum number of tokens in a speech sample for a particular phonological context.

Much variationist research on sound change has focused on vowels. Since vowels are produced along a continuum, intra and inter-speaker variation in their production can often be difficult to discern impressionistically. This has made audio recording technology indispensable to the study of vowel variation and change. As a consequence, the study of vowel variation has benefitted tremendously from advances in recording technology and in computational methods. Another reason for the huge emphasis on vowels is the heavy focus on English. Labov says, "among the languages of the world, English is one of the few that have developed or maintained complex systems of more than 10 vowels. This is not the most promising base for developing general principles of linguistic change" (1994:30). Large-scale acoustic studies of vowel variation such as (Labov, Yaeger, and Steiner 1972) and the *Atlas of North American English* (Labov, Ash, and Boberg 2006) have shown that the key features distinguishing dialects of English are their vowel systems. Vowel variation and change has, thus, played a central role in theoretical developments within the variationist paradigm. Yet, it remains to be seen the extent to which variationist principles of vowel variation and change apply to non-Indo-European (or even non-Germanic) languages that also have large vowel systems like Cantonese.

2.3.2 Structured Heterogeneity Introduced

In an article entitled *Empirical Foundations for a Theory of Language Change*, Weinreich, Labov, and Herzog (1968) lay out what would soon become foundational principles of the variationist approach to language change. One important innovation introduced was the concept of "structured heterogeneity." This concept introduced an attempt to integrate a model of society with a model of language change by recognizing how both society and language may be inter-related to each other. Related to "structured heterogeneity" is the concept of the "speech community," which has since become a fundamental yet debatable concept in the sub-field. The model introduced by Weinreich, Labov, and Herzog (1968) and more finely developed specifically for sound change in Labov (2007) involved a re-framing of the Neogrammarian Controversy as one involving the speech community as a unit of analysis. This led to a shift from language internal vs language external to speech community internal vs. speech community external.

Weinreich, Labov, and Herzog (1968) contextualized their work as an attempt to resolve the Saussurean paradox between linguistic structure and linguistic change. While 19th century linguists were focused largely on diachronic change based on written evidence, 20th century linguists became increasingly interested in the synchronic structure of spoken language. The work of Saussure was pivotal in shifting the concerns of linguistics from the diachronic to the synchronic. This shift was motivated by the Saussurean dichotomy between *langue* (the locus of linguistic structure) and *parole* (speech, the locus of change). This dichotomy posed problems for historical linguists interested in change. As they state, “For the more linguists became impressed with the existence of structure of language, and the more they bolstered this observation with deductive arguments about the functional advantages of structure, the more mysterious became the transition of a language from state to state” (Weinreich, Labov, and Herzog 1968:100).

For example, descriptive linguists of the early 20th century working within Bloomfield’s framework were focused on documenting diverse languages in terms of their synchronic structure. This focus on structure ignored variation within a speech community. Instead of attempting to deal with variation, Bloomfield advocated for abstracting away from the speech of individuals and proposed that the goal of linguistic description should be language at the community-level. This also applied to change. Bloomfield says that the Neogrammarians believed that sound change is a correct assumption “because it alone has enabled linguists to find order in the factual data, and because it alone has led to a plausible formulation of other factors of linguistic change." (Bloomfield 1933:364). This illustrates what Weinreich, Labov, and Herzog (1968) meant by bolstering the advantages of structure with deductive arguments. Thus, following Bloomfield's argument, the assumption that sound change is real accounts for why many languages have systematic sound correspondences in basic vocabulary. Sound change was a concept that showed how all of these languages could be traced to a common ancestor language. It provided a plausible explanation to the diversification of modern languages while simultaneously assuming that all of these languages share a common inheritance. The assumption of regular sound change is an essential assumption in the Comparative Method, which has since been applied not only to Indo-European languages but also to the reconstruction of linguistic history in other parts of the world. This includes languages that have lacked written records. If anything, the success of the Comparative Method beyond Indo-European strengthened the hypothesis that sound change is a real phenomenon that is consistently regular. Yet, the support for sound change was based primarily on deduction.

Deduction is also what characterizes Chomsky's Generative Model of language, which was quickly gaining popularity at the time of Weinreich, Labov, and Herzog (1968). For Weinreich et al, however, the Chomskyan model of "language as a homogeneous object … represent[ed] a backward step" (1968:100). Chomsky reframed Saussure's *langue* vs. *parole* dichotomy as a dichotomy between *competence* vs. *performance*. This placed language in direct opposition to the social world. It was in this theoretical milieu that Weinreich, Labov, and Herzog (1968) saw a major problem in the study of change. They asked, "if a language has to be structured in order to function efficiently, how do people continue to talk while language changes, that is, while it passes through periods of lessened systematicity?" (Weinreich, Labov, and Herzog 1968:100). The solution proposed was one that recognized both language and society as having structure and in showing how the two relate to each other. This was a critique of Chomsky's approach and a return to Saussure and Bloomfield's treatment of both *langue* and *parole* as social but went further than Saussure and Bloomfield did in attempting to explain how social structure relates to linguistic structure and change.

Weinreich et al state that "The key to a rational conception of language change -- indeed, of language itself is the possibility of describing orderly differentiation in a language serving a community" (1968:101). One example of orderly differentiation was Labov's work on rhoticity in New York City. Labov showed how both inter-speaker and intra-speaker variation in the pronunciation of post-vocalic /r/ are connected to each other by showing how both kinds of variation are tied to a socio-economic hierarchy. He examined speakers representing different socio-economic backgrounds as evidenced by surveying employees at three different department stores each representing different socio-economic classes. He showed that post-vocalic /r/ was becoming a prestige variant that was spreading across New York City. The way it spread is reflected in both inter-speaker and intra-speaker variation. At the community-level, there is class-based stratification in terms of who uses the prestige form the most often. The ones most responsible for propagating the spread of the prestige form, however, are not the ones from the highest end of the socio-economic hierarchy, but rather those from the lower-middle class. While the upper-middle class may be the ones who use the prestige variant the most often in casual speech, this is not the case in more formal speech styles. In formal speech styles, lower middle-class speakers surpass upper middle class speakers in usage of prestige variants.

Labov interprets these findings as evidence that there exists a New York City speech community characterized by orderly differentiation as well as a set of norms. The unified set of norms appear at the level of the individual where we see speakers from all socio-economic backgrounds pronouncing the prestige form most often in formal contexts and least often in less formal contexts. Although the exact percentages vary by socio-economic status, all speakers exhibit the same pattern of higher prestige variant usage in more formal contexts. This pattern of orderly differentiation and its link between language use and socio-economic class in New York City illustrate what Weinreich, Labov, and Herzog (1968) meant by "structured heterogeneity". This, they argued, accounted for how social structure and linguistic structure could be connected to each other. This was a relationship that Labov's predecessors were unable to empirically show.

2.3.3 Multilingualism as an example of structured heterogeneity

The main topic of this dissertation is sound change within a heritage language community. This is a specific type of multilingual community. For Weinreich, Labov, and Herzog (1968), multilingual communities are no different from monolingual communities in being characterized by structured heterogeneity. This is the case even if the languages present in a community are genetically and typologically distinct languages. What is most important is the social relationship between the languages present within a single speech community rather than genetic or typological distance of the languages involved. To support their point, they discuss Gumperz's work on long-term intimate contact between Marathi and Kannada in Kupwar, India. They note that even though Marathi and Kannada are so sharply different from each other in terms of surface structure, the two languages "have in fact become so similar that mechanical translation appears to be quite feasible through a simple dictionary look-up procedure" (Weinreich, Labov, and Herzog 1968:158). They argue that the Kupwar case illustrates an example of coexisting systems that are used within specific social contexts. They also point out that these coexisting systems exist in many other communities including ones in which the different systems involved are more closely related dialects. They suggest that the way to conceptualize multilingualism is to recognize that the use of a particular system in all communities is governed by social contextual factors. Thus, code-switching on the individual speaker level is not random but highly constrained by social norms present in particular speech communities. These norms are part of the heterogeneous structures that they theorize.

Weinreich, Labov, and Herzog (1968) downplay a distinction between dialects and languages. They state that, "in principle, there is no difference between the problems of transference between two closely related dialects and between two distantly related languages" (Weinreich, Labov, and Herzog 1968:158). For example, the New York City /r/ study illustrated how individual speakers can switch from not producing /r/ to producing /r/ given the right social context. In this case, the varieties involved are arguably similar enough to be considered dialects, but Weinreich et al. state that "nativelike command of heterogeneous structures is not a matter of multidialectalism or 'mere' performance, but is part of unilingual linguistic competence" (1968:101). They interpret the use of various registers in a community as part of what constitutes "competence" as a member of a particular speech community. They argue that this is also the case in Kupwar, which involves genetically distinct languages that have undergone structural convergence through social norms that have brought these languages closer together. Weinreich, Labov, and Herzog (1968) see the same processes of linguistic transfer in Kupwar as they see in New York City.

The concept of the structurally heterogeneous speech community, thus, includes formulation of a very specific relationship with the different varieties of speech (whether this means different dialects or different languages) spoken within a speech community. This particular perspective of variation has essentially remained unchanged in Labov's more recent work. For example, in describing results of research on New York's Lower East Side, Labov says that "the striking regularity" observed "demonstrates that variation in the urban speech community is not the chaotic result of dialect mixture, but a highly constrained pattern that closely determines the linguistic behavior of each speaker" (2001:80). In discussing the Transmission and Diffusion model, Labov says "that there is no substantive difference" (2007:346) between language and dialect. This is the case even though the bulk of variationist sociolinguistics research on vowels has focused on cases that arguably involve dialect contact rather than language contact. Although the difference between dialect and language contact can be a fuzzy distinction in some cases, the case of contact between Cantonese and English in Toronto is not a fuzzy case. Cantonese and English belong to two genetically and typologically distinct language families. Furthermore, as I discuss later on (especially in Section 7.1), speakers sometimes consciously avoid phonological convergence.

2.4 THE TRANSMISSION AND DIFFUSION MODEL

The Transmission and Diffusion Model (Labov 2007), which is the central theoretical focus of this dissertation, is an attempt to reconcile the debate between the Family Tree Model and the Wave Model (including lexical diffusion) in accounting for language change by discussing two mechanisms for change that are the result of two different types of language learning. The first is *transmission*, which involves the "unbroken sequence of native-language acquisition by children" (Labov 2007:346) while the second one is *diffusion*, which involves the transfer of features from one speech community to another via adult language learning. These two types of changes are related to the critical period concept in language acquisition research and are argued to have implications for language change at the level of the speech community. Labov argues that "transmission is the fundamental mechanism by which linguistic diversity is created and maintained" while "diffusion is a secondary process, of a very different character" (Labov 2007:347). His rationale for this important distinction is influenced by Scovel's extreme view that a "native-like accent" for adult learners is "impossible" (2000:217). Thus, he assumes that while children are successful at learning the phonology of the language spoken in their speech community, adults fail at learning the phonology of the dialect or language of another speech community.

Transmission begins with child language acquisition when adults pass on their linguistic system to their children. The transmission process includes transmission of both lexicon and structure. This, Labov argues, accounts for the large degree of inter-generational continuity observed in the historical development of languages. When children do end up modifying the adult system, they modify the system in a highly structured way. Thus, any change initiated by children is change due to general cognitive principles and constraints. In other words, children initiate change due to internal motivation. Adults on the other hand have more limited language learning abilities. They introduce externally motivated change when they learn another language or dialect. "The adults who are the borrowing agents," however, "do not faithfully reproduce the structural patterns of the system they are borrowing from" (Labov 2007:383).

Labov (2007) illustrates this in fine detail in his discussion of short /æ/ patterns across several dialects of American English including Cincinnati, New Orleans, Chicago, and Saint Louis. He ties the development of the short /æ/ patterns (such as which words have lax vs. tense /æ/) in these different cities to historic migration patterns involving speakers from New York City who settled in places across the expanding territory of the US in the 19th century. The original New York City system involves a complicated set of grammatical and phonetic constraints in addition to lexical exceptions that account for which lexical items have a tense /æ/ and which have a lax /æ/. /æ/ is tensed in closed syllables, before some front nasal clusters, and all front voiceless fricatives. A newer system developed in the 20th century involves expansion of the tensing environments to also include all front nasal codas, and all voiced stops. The short /æ/ system also involves a function word constraint. Function words such as <can> have lax /æ/ while <can> as a content word has tense /æ/. There are also lexical exceptions such as tense /æ/ in <avenue>, but lax /æ/ in many other words with a similar phonetic environment such as <average>, <savage>, and <gavel>. Some exceptions involve words that are typically learned later in life such as <carafe>, which would be lax even though the /æ/ precedes a fricative, an environment that otherwise conditions tensing.

Labov (2007) argues that only children can master these complex set of rules and constraints. When settlers from New York City migrated to places such as New Orleans and Cincinnati in the 19th century, the inherited New York City pattern in these cities became modified through *diffusion* and then *transmission*. So in other words, the first step was dialect contact among adults who introduced new pronunciations of some words with short /æ/. This would be *diffusion* and since adults initiate diffusion, what adults borrow is lexicon rather than a new structural pattern. In the *transmission* process from adult to children, children will create a new pattern that is more structured than the irregular pattern introduced by adults through *diffusion*. This new pattern is then passed on to subsequent generations. This is what leads to long-term change.

These two mechanisms also correspond to the distinction Labov made in earlier work between "change from above" and "change from below". The source of "change from above", which is arguably a form of dialect contact (Guy 1990:51), is a socially prestigious group with a prestigious pronunciation pattern. This pattern is irregular and unsystematic in contrast to the vernacular language acquired by children. One example of "change from above" is the diffusion of post-vocalic /r/ in New York City during the 1960s when (Labov 1972) conducted his pioneering study of /r/ usage across different social classes. Change from below, however, is regular and often unconscious. The fact that acoustic analyses have shown many cases of change in vowel systems even though speakers generally do not notice such changes illustrates how changes in vowels are typically changes from below. Though adults may pick up new vowel pronunciations for some words unconsciously, the new patterns they introduce are not regular according to Labov (2007). Thus, in formulating the distinction between change from above and change from below, Labov (2007) also formulates a relationship involving this dichotomy and specific phonetic features.

According to Labov (2007), the long-term consequence of these two mechanisms is reflected in the types of changes depicted by both the Family Tree Model and the Wave Model. Thus, *transmission* is what results in internally motivated change and is what makes it possible to reconstruct linguistic history based on the assumption of exceptionless sound change. *Diffusion*, on the other hand, is what accounts for the exceptions to sound change. Diffusion is also what is accounted for in the Wave Model. In addition, Labov (2007) argues for a strong constraint against structural diffusion. The implication of this constraint is that structural changes develop only through transmission. When structural borrowing does occur, it is always initiated by the borrowing of loan words. Chain shifts, for example, may be diffused or transmitted. Structural changes such as vowel splits, however, do not spread through contact.

Labov (2007), however, does mention that mergers may be an exception to his proposed constraint against structural borrowing because mergers lead to a loss in contrast. Since it is easier for adults to learn fewer phonological contrasts, it thus follows that mergers are not subject to the same constraint according to Labov (2007). The example he cites is Herold's (1990; 1997) research on the development of the low-back merger in the historic anthracite mining region of Eastern Pennsylvania. Herold (1997) found a correlation between a town's historic economic dependence on mining and whether or not speakers from that town have the low-back merger. She attributed the development of the low-back merger in these towns to the large influex of Eastern European immigrants who settled in these towns. Many of the languages spoken by these immigrants have five-vowel systems (including Polish and other Slavic languages as well as Italian) and hence a lack of a low-back distinction. The merger would have then been introduced through sturctural influence from these heritage languages.

Thus, to summarize, Labov (2007) acknowledges two different types of processes that are initiated by different agents (children vs. adults) that lead to different patterns. One pattern is regular, internally-motivated, and reflected by the Family Tree Model while the other is irregular, externally-motivated, and reflected by the Wave Model. Direct structural innovation is possible in the former but not in the latter except for simplifying changes such as vowel mergers.

2.5 CRITIQUES OF LABOVIAN APPROACHES

Kiparsky has recently praised the Labovian framework for providing “decisive evidence for neogrammarian across-the-board sound change” (Kiparsky 2016:465) and evidence against lexical diffusion as a primary mechanism of change. Labov’s research on sound change has also been very influential among historical linguists (Garrett 2014). Ironically, much of the critique of Labov’s framework has come from within sociolinguistics, the subfield he inadvertently created. In this section, I discuss several major issues in Labov’s framework (which I will refer to synonymously as a “First Wave” approach) that have been critiqued by sociolinguists. I begin by discussing the major issues that have been the focus of Third Wave studies (Eckert 2012) in Section 2.5.1. I then discuss how heritage language use poses problems for two concepts that have been foundational in the Labovian paradigm: the speech community and the vernacular (Section 2.5.2).

2.5.1 The Third Wave critique of macro-sociological variables

Eckert (2012) describes “three waves” of variationist research. These three waves are defined primarily in terms of theoretical orientation and approach to investigating language variation and its relationship to the social world rather than in terms of chronological order. Much of the research that led to the Transmission & Diffusion Model would fall under the “first wave”, which Eckert (2012) describes as focusing on macro-sociological variables and how they correlate with use of different linguistic features. The second wave refers to research that takes an ethnographic approach. In doing so, second wave studies focus on the local meaning of the macro-sociological categories that are the focus of first wave studies. Like in first wave research, however, the second wave treats identity as category affiliation, whether the category is pre-defined by the researcher or by the local community. The third wave differs from the other two waves in flipping the relationship between language and society. In other words, rather than treating language variation as a reflection of social category membership, third wave studies focus on linguistic practice and how individual speakers create social identity through the use of different styles. This approach places much more focus on individual speaker agency and leads to a more fluid and dynamic view of identity and social meaning. The social unit of analysis is also shifted from the speech community to communities of practice.

One illustration of the third wave approach to sound change is Eckert's study of the pre-nasal split of /æ/ in two Northern California elementary schools (2008). On a broad geographical scale, the nasal split of /æ/ is a feature of white Anglo California speech. Under a first wave approach, the social significance of the split would focus on the correlation of the split with the speech of white Anglo speakers and the absence of the split with the speech of other ethnic groups. Specific instances of speakers of other ethnic groups adopting this feature would be interpreted as assimilation to the regional norm. By focusing on ethnographic research in two elementary schools with different ethnic demographics, Eckert (2008) shows how the social significance of this variable can take on different meanings on the micro-scale and how these different meanings arise from the different social contexts created in the different school environments. For instance, at Fields (a pseudonym) Elementary School, where white Anglo students are the numerical majority, it is the regional variety of white Anglo speech that has symbolic capital. Thus, not surprisingly, most of the children at this school have the pre-nasal /æ/ split that is characteristic of white Anglo California speech. At Steps (another pseudonym)

Elementary School, however, where Latino children are the majority, Chicano English has symbolic capital.

Although the lack of a pre-nasal split can be described as a substrate feature influenced by Spanish, which lacks this feature, Eckert argues that it would be oversimplistic to treat the lack of a pre-nasal split as a feature that only indexes ethnicity. She argues instead that the meaning of this variable is underspecified. This makes it possible for the presence or absence of the split to index different meanings depending on the school context and on specific interactional moments. For instance, even though there are white Anglo students at Steps who produce the pre-nasal split, the split pattern is weaker than the pattern at Fields. At Fields, it is not ethnicity alone that predicts which speakers have the strongest split. It is the girls who participate in the popular crowd who lead in the use of this feature regardless of ethnicity. For example, the speaker who has the second strongest split is in fact a Latina girl, who lives in the same neighborhood where many Latino students who attend Steps also live. She was sent to Fields because her parents wanted her to attend a safer school, but she has clearly developed a pattern that is more typical of white Anglo speakers in spite of her family background and place of residence.

Eckert (2008) describes her argument as one that appears subtle to the sociolinguist seeking larger patterns. She argues that the explanation for variation lies not in social categories (or even in the speech community), as has been the focus of first wave studies, but in social practice. In other words, the reason that some individual speakers produce the split while others do not lies in how individual speakers construct their identities in interaction with other speakers in their social environment. The use or absence of this variable could still correlate with ethnic group membership on a broader geographical scale, but focusing exclusively on these broad patterns makes it more difficult to see how ethnicity and other broad social categories are constructed in micro-level social interaction. This leads to a less deterministic view of social categories. This makes it possible to account for speakers who do not use the variants expected based on their demographic group membership. The use of these variants is more closely related to participation (and hence as a practice) in the popular crowds at each individual school than to membership in an ethnic group (or structural position within a speech community).

2.5.2 The problem of heritage speakers

An important difference between a first wave and a third wave explanation is different analytical units of social organization. Under the first wave approach, it is the speech community while the second and third waves have focused on "communities of practice". This shift from "speech community" to "communities of practice" parallels the shift from fixed social membership categories to how speakers construct these categories through linguistic practice. It also leads to a different way of conceptualizing multilingualism. Although multilingualism per se has not been a major focus of third wave variationist research (with some exceptions to be discussed shortly), many researchers who focus on the sociological aspects of multilingualism have adopted the communities of practice concept.

Romaine, for example, says that "we all belong to many communities and sub-communities, defined in terms such as social class, ethnicity, nationality and religion" (2012). This is a much broader view of "community" than what is entailed in the speech community. This is also a view that recognizes the possibility of membership in multiple communities that may be defined at least partially by the use of particular forms of speech. She further elaborates by describing bilingual communities in ethnic neighborhoods throughout the world such as Little Haiti in Miami, Little Italy in Boston, and the Chinatowns of various cities. In these communities, speakers belong not only to broader communities that include speakers of their heritage languages but also to the global community of English speakers even if English proficiency levels are variable.

Romaine says that, “What is crucial, then, to most definitions of community is the sense of perceived solidarity and interaction based on reference to a particular language and the relationships among people who identify themselves as members of that community. In this sense … they constitute ‘communities of practice’ (2012).” The communities of practice model, thus, can describe the maintenance and continued development of sociolinguistic norms across geographic space as illustrated in the examples of ethnic communities in cities dominated by English. It can also account for speakers who live in the same neighborhood but do not use all of the languages they speak to communicate with each other. For instance, a multilingual Hong Kong immigrant in Toronto would use English to communicate with a monolingual Torontonian neighbor. In this sense, there is a community formed through the common use of English. The Hong Kong immigrant, however, might maintain ties with Hong Kong and continue to interact with Cantonese speakers in Hong Kong. This would be how the Hong Kong immigrant would continue to be part of the broader global Cantonese speech community as well as the local one in Toronto.

It may in fact be questionable as to whether or not the speech community model was even intended for the type of sociolinguistic context involving heritage language speakers. As Labov says, “such a clear dichotomy between transmission and diffusion is dependent on the concept of a speech community with well-defined limits, a common structural base, and a unified set of sociolinguistic norms (2007:347)." In the case of a heritage language spoken in a diasporic context, it seems doubtful that these assumptions are met. First of all, Romaine's discussion of communities of practice that cut across geographical space does not seem consistent with the idea of a clearly delineated speech community. English and Cantonese also involve very different structural bases since they are genetically and typologically distinct languages. In the Kupwar case, a common structural base developed through structural convergence, but this presumably took multiple generations to develop. Furthermore, there is no evidence that there are many non-Chinese heritage English speakers in Toronto who learn Cantonese. While Cantonese may be converging with English due to bilingual speakers who speak both languages, most Toronto English speakers do not learn to speak Cantonese since Cantonese is primarily an in-group language. Cantonese also lacks official language status. Hence, the Toronto case lacks the multi-directional influence found in the Kupwar case.

Also important in the speech community concept as well as in the Transmission & Diffusion Model is the concept of the vernacular. This has also been the subject of much discussion in third wave critiques (cf. Bucholtz 2003; Eckert 2003; Eckert 2008; Eckert 2012; Coupland 2016 among others). Under first wave approaches, style is treated as a continuum based on attention paid to speech with the vernacular on one end and formal speech on the other. The assumption in first wave approaches is that the vernacular is the language that a speaker first acquires from their parents and is hence assumed to be the most systematic and least affected by social correction of language use. Under the third wave approach, however, the individual development of stylistic repertoires is an ongoing process throughout an entire person's life. Although some third wave scholars have recognized age effects (either as gradual or as a critical period), the focus of many third wave studies has been on how individual speakers make use of different linguistic resources available to them in social interaction rather than on testing the limitations of the critical period concept. For example, Eckert says that "it is not clear … how much the age constraints on acquisition, and perhaps particularly of dialect features, are due to cognitive limitations and how much to the social conditions under which we learn and use language" (Eckert 2003:395). Furthermore, it is also not clear how stable the phonology of a vernacular is throughout a person's lifetime if the vernacular is a heritage language.

Heritage speakers who acquired two languages at an early age in a societal context in which their second language is the locally dominant language are a group of speakers that pose problems to both the use of critical period concept in variationist sociolinguistics and to the vernacular. All of the speakers included in this dissertation learned Cantonese as their first language and English as their second language. GEN 2 speakers also acquired both languages at an early age. Cantonese would clearly be their vernacular based on order of acquisition. Following first wave assumptions, we might expect their Cantonese to be more systematic than their English. We might even expect Cantonese influence on their English but not the other way around because the vernacular is assumed to be immutable according to the logic in First Wave studies. As I show in the next sub-section, however, many variationist studies suggest that heritage language (L1) to dominant language (L2) influence is not always the case. Furthermore, the role of relative age of acquisition vs. linguistic dominance has been a debatable topic among researchers in both contact linguistics and in psycholinguistic approaches to bilingualism as I will discuss in Chapter 3.

2.6 RESEARCH GAPS IN VARIATIONIST APPROACHES

The central theoretical problem of this dissertation is whether or not contact-induced phonological change in the vowel system of a heritage language can develop within two generations in native vocabulary. In Labov's (2007) TD Model, contact-induced phonological changes are theorized as an adult initiated process mediated by lexical borrowing. What is not recognized under this model is the possibility of children acquiring two or more languages at an early age and introducing contact-induced phonological changes from the dominant language to the heritage language. Very few empirical studies using variationist approaches have investigated this issue.

Most variationist studies that do investigate the role of children (or teenagers) in initiating contact-induced phonological changes have focused on influence from minority or heritage languages on the societally dominant language rather than the other way around. This is a topic that has become an increasingly major topic of variationist research. Many of these studies, however, have been framed in terms of ethnic group participation in regional sound changes rather than in terms of heritage language influence on the dominant language. In fact, individual identity has generally been shown to have better explanatory value than influence from the actual phonological structures of heritage languages.

Eckert's (2008) study discussed in the previous section would be one example of such research. Fought (2003) has also examined vowel pronunciation patterns in the English spoken by heritage Spanish speakers and has also shown how various aspects of identity (such as gang vs. Non-gang affiliation) are related to these pronunciation patterns. In Cajun English, Dubois and Horvath (1999) discuss a v-shaped pattern in several phonetic variables across three generations of speakers. The first generation was the first English-French bilingual generation and showed the most phonetic influence from Cajun French. The second generation grew up during a time of heavy stigmatization against the use of French and consequently adopted more Southern English variants. The third generation, however, grew up during the time of the Cajun Renaissance, a period during which Cajun identity became a source of local pride. As a consequence, some of the Cajun English features used by the first generation became "recycled", but it is specifically third generation men who lead in the use of Cajun features because of an indexical link between traditional Cajun cultural practices and Cajun male identity.

What these aforementioned studies make clear is that heritage languages do not have a deterministic force on the pronunciation patterns of the dominant language spoken by bilingual speakers. This is also true when the heritage language undergoes language loss and a new ethnic variety develops as in the Cajun English example. Another example illustrating this point in rich ethnographic detail is Bucholtz's (2009) study of two Laotian refugee teenage girls in a Northern California high school. These two girls had very similar backgrounds before entering high school. In high school, however, one of them became involved with gangs while the other one avoided gang culture. While the former had features approaching African-American Vernacular English, the other had features approaching the local dominant white Anglo variety of English. The former, for example, did not pronounce post-vocalic /r/, while the latter did. One's individual identity defined in terms of the communities of practice in which they participate, had a better explanatory role in the features that appear in the English spoken by these heritage speakers than the features actually present in their heritage languages.

Some research on ethnic group participation in regional sound change has also included Cantonese heritage speakers. For example, Wong and Hall-Lew (2014) compared Chinese-

Americans (including many Cantonese heritage speakers) in both New York City and San Francisco. While younger Chinese-Americans in San Francisco were shown to lead in the low-back merger, those in New York City have all maintained a distinction. These patterns conform with the regional patterns of other groups in these regions. The presence of a distinction between two similar vowels in Cantonese did not prevent younger Cantonese heritage speakers from merging these two vowels in the English that they speak in San Francisco. Research on the English spoken by various ethnic groups in the Toronto area also shows Cantonese heritage speakers sharing regional features, such as the Canadian Vowel Shift, with speakers of other ethnic groups (Hoffman 2010; Hoffman and Walker 2010).

Another strand of variationist research on the influence of heritage languages on the dominant language has focused on the development of what are referred to as 'multiethnolects'. These studies have taken place primarily in urban European contexts. One example is Multicultural London English (MLE), which Cheshire et al. (2011) have described as a new variety of English that has arisen through group second language acquisition of English. The sociolinguistic context that makes this possible is the presence of a large immigrant community coming from many different parts of the world and speaking many different heritage languages. Unlike many of the cases researched in North America, this strand of research has focused on the development of new dialects created by multiple ethnic groups. Similar research has also been conducted on an emerging form of Australian English (Kiesling 2005). Like North American studies of the speech of individual ethnic groups, however, such research has focused on varieties of the locally dominant language rather than potential influence of the dominant language on heritage languages.

The lack of variationist research addressing influence going in the other direction is part of what has motivated the Heritage Language Variation and Change in Toronto (HLVC) Project (Nagy 2011). This project has involved collection of sociolinguistic interviews from speakers of various heritage languages. Studies conducted to date using this data have involved a wide variety of variables including Pro-drop, VOT, noun/verb paradigms, and final devoicing[3]. The only language for which vowel analysis has been conducted is Cantonese. Published studies of Toronto Heritage Cantonese have shown overall maintenance of phonological contrasts for at least four of the contrastive monophthongs in Homeland Cantonese (Tse 2016a; 2016b). In Tse (2016b), I showed that an allophonic distinction between [ɪ], which occurs only before velar consonants, and [i], which occurs elsewhere, has been maintained although there is evidence of low-level phonetic changes in these vowels initiated by GEN 2 women. In the same paper, I showed that a similar allophonic distinction between [ʊ], which also occurs only before velar consonants, and [u], which occurs elsewhere, is also a feature that is unchanged in GEN 2 speech. These findings show support for treating the Cantonese of GEN 2 speakers as a vernacular (in the Labovian sense) that has been faithfully transmitted from the previous generation. Results from Tse (2016a), however, show evidence that leads one to question the immutability of the vernacular. This study showed evidence of GEN 2 speakers innovating allophonic splits for two different vowel categories. GEN 2 men lead in splitting /ɔ/, but no sex-based difference was identified in the split of /ɛ/. It was unclear whether these changes are internally motivated, due to contact-induced change, or due to internally motivated changes that follow contact-induced change.

[3] See http://projects.chass.utoronto.ca/ngn/HLVC/1_7_variables.php for extensive list of references.

Both of these studies, however, were based on a limited set of data. In this dissertation, I expand on the analyses developed in these two papers by using a larger set of data and a larger set of speakers. The dissertation includes data from Homeland speakers to address whether or not any changes identified among GEN 2 speakers are also changes found in Homeland speech. This is especially important for addressing whether or not the allophonic splits identified in Tse (2016a) are due to contact-induced change or whether they are internally motivated by pre-existing tendencies in Cantonese such as preservation of contrasts in pre-velar environment.

Showing that these are contact-induced changes would present a problem with the transmission vs. diffusion dichotomy since (Labov 2007; 2011) argues that structural change cannot be diffused with the exception of mergers. Finding evidence of allophonic splits diffused through contact with Toronto English would, thus, present a problem to this dichotomy since an allophonic split would be a structural change. Identifying contact-induced splits would also address a paradox that (Labov 1994:331) observed when he noted that the majority of structural changes identified in variationist research on sound change has involved mergers rather than splits. He says that if this is representative of all possible changes, then it would follow that languages would be simplifying their phoneme inventories over time, but that is clearly not the case. In Tse (2016a), I argued that the reason for this paradox is the lack of research on sociolinguistic settings in which such structural changes would be more likely to occur. Identifying allophonic splits in this dissertation would provide further evidence supporting this claim.

The next chapter will focus on debates about the role of age of acquisition vs. linguistic dominance in theories of contact-induced change. This will motivate the hypotheses that will be presented in Chapter 5. This is a question that has not been a major focus of variationist research.

This body of research has had very little direct engagement with variationist research on sound change. For this reason, discussion of these debates are presented in a completely different chapter.

2.7 CHAPTER SUMMARY

This chapter began with a review of the Neogrammarian Controversy. Is sound change an exceptionless process or is it the case that “every word has its own history”? After discussing the two sides of this controversy, I showed how new ideas and technological advances from the 1960s onward transformed the original debate. The tools of laboratory phonetics made it possible to observe sound change in progress. The Lexical Diffusion Model (Wang 1969; 1979) introduced a modified form of the Wave Model by explaining how some changes may appear exceptionless while others appear to have many exceptions. I also discussed the theoretical foundations of variationist sociolinguistics by reviewing important ideas introduced in Weinreich, Labov, and Herzog (1968). I showed how this eventually led to the development of the Transmission and Diffusion Model. The key differences between transmission and diffusion are summarized in Table 3. After discussing the Transmission and Diffusion Model, I presented critiques of not only the model itself but of the research framework developed by First Wave variationist studies. This critique also discussed alternatives such as the communities of practice concept. I also discussed how speakers of heritage languages who subsequently become dominant in another language pose problems with both the speech community concept and the vernacular, which are both concepts central to the theoretical basis of the Transmission and

Diffusion Model. After this discussion, I presented a general overview of variationist research that has addressed related topics as a way of identifying gaps in research knowledge.

Table 3. Transmission vs. Diffusion

	Transmission	**Diffusion**
Acquisition Mechanism	L1 child acquisition	Adult L2 acquisition
Speech community relationship	Within a speech community (internal motivation)	Across speech communities (external motivation)
Style	Minimum attention paid to speech (vernacular)	More formal, more attention paid to speech
importance to sound change	"the fundamental mechanism by which linguistic diversity is created and maintained" (Labov 2007:347)	"a secondary process of a very different character"
Source of change	change due to general cognitive principles and constraints	unsystematic changes due to social reasons or contact
Social implications	Change from below	Change from above
Social Awareness	Low	High
Linguistic unit	phoneme	Word
Possible phonological outcomes	Chain shifts, splits, mergers	Chain shifts, mergers
What it explains	Neogrammarian sound change	Lexical Diffusion
Model	Family Tree Model	Wave Model

The key theoretical question of this dissertation is whether or not the Transmission and Diffusion Model can be applied to studying sound change in an L1 that is initiated by speakers who subsequently become linguistically dominant in an L2 during childhood. This is a question that has not been addressed in previous studies of vowel variation and change in progress. This dissertation is part of the first major survey of what kind of variation in vowel production exists in a heritage language and is also the largest variationist study of Cantonese vowels. The central point of the current study is to show how the actual patterns observed reveal problems and limitations with the Transmission and Diffusion Model. I argue that more suitable frameworks are ones proposed by contact linguists as I discuss in the next chapter. While the methodology of this study is more aligned with the first wave approach, the explanation of the results has been influenced by Third Wave ideas. Ultimately I will argue that the Transmission & Diffusion Model is indeed a model unable to account for settings similar to what can develop among a group of heritage speakers. The solution I propose is for variationists to further refine models of contact-induced change that have been developed by other scholars such as Thomason and Kaufman (1988) and (van Coetsem 1988; 2000). These two models will be discussed in the next chapter.

3.0 CONTACT LINGUISTICS AND BILINGUAL PHONETICS/PHONOLOGY

The focus of this chapter is on contact-induced change from the perspective of contact linguistics and heritage language bilingual phonetics and phonology research. I show that research in these areas can offer valuable perspectives for understanding the factors that could potentially influence variation and change in the vowel system of heritage speakers. The two models of contact-induced change introduced in this chapter include Thomason and Kaufman's (1988) Analytical Framework for Contact-induced Change (henceforth "TK Model") and van Coetsem's (2000) General and Unified Model of the Transmission Process in Language Contact (henceforth "VC" Model). The advantage of these two models over Labov's TD Model is that they both consider a broader range of types of contact settings including those involving heritage language bilingualism. They have also focused on the specific conditions that facilitate different types of contact-induced change such as lexical borrowing and structural influence. Despite consideration of a broader range of possible sociolinguistic settings, skepticism against explanations based on contact-induced change persist partly due to methodological reasons.

Where these two models differ from each other is in their treatment of the underlying mechanisms behind contact-induced change. The TK Model is based on sociolinguistic continuity while the VC Model is based on individual speaker linguistic dominance. A variationist study of vowel variation and change within a community of bilingual heritage speakers would be an empirical study well-suited to address the advantages and disadvantages of each model since inter-generational change in linguistic dominance is one of the characteristics of the community examined in this dissertation. The type of vowel changes (splits, mergers, or shifts) that actually develop among GEN 2 Cantonese speakers is the central focus of this dissertation. If there are changes, can they be shown to be contact-induced changes? If so, what are the implications for understanding sound change?

In Section 3.1, I introduce both the TK and VC Models. I then compare and contrast these two models along with the TD Model in Section 3.2. One important difference is that the TD and TK Models are based at least partly on age of acquisition while the VC Model is based on individual linguistic dominance. A model based on linguistic dominance seems more appropriate for the study of heritage language bilingual speech, but as I discuss in Section 3.3, studies of early bilingual speakers, including heritage speakers, have generally shown a *duality of patterning* (Hockett 1958) characterized on the one hand by phonemic systems resistant to change but on the other hand, by grammatical systems open to change. Some of the studies discussed in this section also show low-level phonetic change. What is not clear from experimental studies is whether or not such low-level phonetic changes are indicative of changes in progress. In Section 3.4, I discuss how the variationist approach can address several problems in studies of contact-induced change and bilingual phonetics and phonology. The study of heritage language vowel variation and change, thus, offers an opportunity to bring variationist sociolinguistics research in dialogue with research in contact linguistics and heritage language bilingual phonetics and phonology. This also brings a psycholinguistic perspective (as in the VC Model) together with a sociolinguistic and sociohistorical perspective (as in the TD and TK

Models). Thus, how is sound change actuated (psycholinguistic perspective) and how is it propagated (sociolinguistic perspective)?

3.1 TWO MODELS OF CONTACT-INDUCED CHANGE

3.1.1 The Thomason & Kaufman (TK) Model

The first model I discuss is Thomason and Kaufman's (1988) Analytical Framework for Contact-induced Change. This framework arose as a critique of the privilege placed on internal motivation over external motivation in historical linguistics. This point is summarized in a widely cited quote stating that, "the history of a language is a function of the history of its speakers and not an independent phenomenon that can be thoroughly studied without reference to the social context in which it is embedded" (Thomason and Kaufman 1988:4). The TK Model takes a socio-historical perspective with the community as the level of analysis and a distinction between sociolinguistic continuity (language maintenance) and discontinuity (language shift) in the group-level use of specific languages as the primary basis for the distinction between contact mechanisms.

The part of the TK model most relevant to the current discussion is shown in Figure 1 below.[4] The TK Model includes two major dimensions: contact mechanism (borrowing under maintenance vs. interference through shift) and intensity of contact. I will henceforth refer to the

[4] Not included are the parts showing the formation of pidgins, creoles, and bilingual mixed languages. Although I recognize potential implications between my study and debates about the formation of these types of languages, discussion of this topic would require going beyond the scope of this dissertation.

processes as "Maintenance" and "Shift" for short. Maintenance in the TK Model refers specifically to cases in which the language of a community is maintained and transmitted to subsequent generations of speakers. Shift, on the other hand, refers to cases in which a community of speakers learns the language of another group. The possible linguistic outcomes that can result from these two types of situations depend on intensity of contact. Intensity has a different meaning depending on whether the situation is one involving Maintenance or Shift.

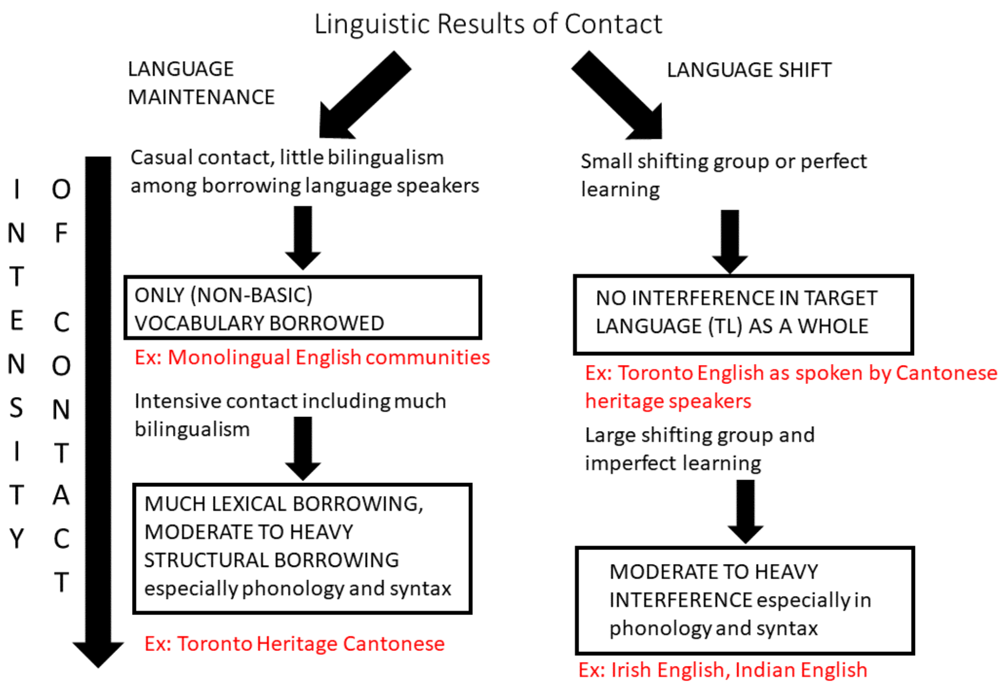

Figure 1. Linguistic Results of Contact (adapted from Thomason and Kaufman 1988)

Under Maintenance, the lowest intensity setting is one that involves monolingualism and casual contact with speakers of another language. This would describe most variationist studies (Nagy and Meyerhoff 2008). Under such a setting, the only type of contact-induced change that is possible is the borrowing of loan words. If speakers in a community do not speak any other language, then it follows that they have no access to the grammatical system of another language. If they lack access to an alternative grammatical system, then direct structural transfer from another language through contact is unlikely, although given time, structural changes could potentially develop from a large set of lexical borrowings.

Higher levels [5] of intensity in contact under Maintenance involve higher levels of bilingualism or multilingualism in a community. Under higher intensity Maintenance situations, structural borrowing especially in phonology and syntax are described as possible. If speakers speak more than one language, then it becomes possible for them to transfer structural patterns from one language to another. Heritage language bilingualism would be considered an example of high-intensity contact under Maintenance. In such a contact situation, the heritage language is maintained but may be influenced by the other language or languages spoken within a community. The other language may be a societally dominant language. For example, Thomason and Kaufman describe Yiddish-speaking immigrants in the US (1988:40). Since English is the societally dominant language, English has an influence on the Yiddish spoken within this community. The influence is strongest on the lexicon and to a moderate extent on the morphosyntax but relatively weak on the phonology (ibid.). Thus, if there is sufficient proficiency in two languages, structural influence from one language to another is possible.

The direction of influence among the different languages spoken within a community are described as involving different contact mechanisms. For example, while the heritage language spoken within a community of heritage speakers would be described as high-intensity

---

[5] Thomason and Kaufman (1988) propose five different levels of intensity in contact under maintenance. For the purpose of this discussion and the rest of the dissertation, however, it would be sufficient to describe these different levels in relative terms (ie lower vs. higher).

Maintenance, the dominant language spoken within the same community would involve a case of Shift. Maintenance differs from Shift in terms of the agents of change and direction of influence. In the case of Maintenance, it is native speakers of a language who initiate changes that have effects on the development of a language. This can also include influence from another language also spoken by these speakers. In a Shift case, it is non-native speakers of a language who initiate changes. Not all cases of Shift, however, lead to language change. This depends on the intensity of contact involved.

A low-intensity case of Shift would involve either a small shifting group or "perfect" learning. Thomason and Kaufman describe urban immigrant groups of European origin in the US as "one of the most typical cases" of language shift without interference on the target language (1988:120). The initial generation that learns to speak the target language (English) would typically speak the target language with phonological and syntactic features influenced by their native language. The English spoken by the Yiddish-speaking immigrants as described above would be an example. After several generations, however, speakers in many such communities lose their heritage language and completely shift to English. The variety of English they speak, however, is typically not much different from the English spoken by other groups of speakers. One recent example illustrating this is the English spoken by Toronto Heritage Cantonese speakers (cf. Hoffman and Walker 2010).

A high intensity case of Shift would be a case in which the structural features of the version of the target language spoken by the initial generation of bilingual speakers are passed down to subsequent generations. This happens under specific social circumstances. For example, Thomason and Kaufman mention immigrant groups who settled in rural areas who lacked access to other groups of speakers of the target language but still ended up learning the target language

(1988:120). Another set of examples that Thomason and Kaufman mention are those involving learning a literary or official language as a second language as in the case of the varieties of English spoken in India and Ireland (1988:129). In these cases, there may be phonological, syntactic, or both phonological and syntactic influence of the native language on these varieties of English. One example illustrating shift-induced phonological change in the development of an English variety is the development of a contrast between dental and retroflex stops (found across many languages spoken across the Indian subcontinent) in Indian English (ibid.).

Another important aspect of the TK Model is the caveat that under the right social circumstances, any of the predictions made in this model can be violated. This also includes lack of change under social circumstances in which massive change would otherwise be expected. The claim that anything is possible is a point that some researchers have critiqued (Labov 2007; Labov 2011; Winford 2003; Sankoff 2013) for being too extreme. Thomason, however, clarifies this point by describing predictions about contact-induced change as "a matter of probabilities, not possibilities" (2001:71). For example, structural borrowing is much more likely to occur in a high-intensity Maintenance context than in a low-intensity Maintenance setting following the TK Model. If we find that structural borrowing does not occur in a high-intensity Maintenance setting such as in Toronto Heritage Cantonese, it would be because of social factors such as language attitudes that discourage structural borrowing.

One example that Thomason (2001:81–82) discusses to support her point involves speakers of Montana Salish, an endangered language. Montana Salish speakers are all bilingual and speak English as their dominant language. This would be a high intensity Maintenance setting, which means that it is a more probable setting for structural borrowing than a low intensity Maintenance setting like a monolingual English speaking community. Montana Salish, however, shows very minimal influence from English. Thomason says that this is because of strong ideological beliefs against mixing languages. Even for new technology items, Montana Salish speakers prefer calquing based on existing Montana Salish vocabulary over borrowing words from English. Thomason also notes that the lack of English influence does not mean that there are linguistic constraints against contact-induced change. To illustrate her point, she discusses an example of a consultant who provided very literal translations of English sentences (Thomason 2001:81–82). These Montana Salish sentences would be, thus, heavily influenced by English morpho-syntax. Although this was based on a misunderstanding of the task, Thomason argues that the fact that an individual speaker could still produce such sentences shows that the reason such sentences are not more common among Montana Salish speakers is because of social rather than cognitive or linguistic constraints. What the Montana Salish example illustrates, according to (Thomason 2001) is the primacy of social factors in accounting for language change or lack of change.

3.1.2 The van Coetsem (VC) Model

The second model of contact-induced change that I present is van Coetsem's (2000) General and Unified Model of the Transmission Process in Language Contact. Unlike the TD and TK Models, the VC Model takes a psycholinguistic perspective with the individual as the primary level of analysis. According to this model, there are three transfer types: RL (Recipient Language) Agentivity [6], SL (Source Language) Agentivity [7] (or "imposition" roughly

[6] Van Coetsem (1988; 2000) uses the terms *RL Agentivity* and *borrowing* synonymously. I will use only the former to refer to van Coetsem's concept to avoid potential confusion with the term *borrowing* since it is a term that

corresponding to TK's Shift), and Neutralization. The psycholinguistic basis of these three transfer types is individual linguistic dominance.

Van Coetsem defines linguistic dominance as "based on the greater proficiency that a speaker has in one language (L1) as compared to another (L2). L1 refers to the language in which the speaker is most proficient, although it is not necessarily his first acquired or native language" (2000:66–67). He says that the reason that linguistic dominance has a strong effect is because of the stability gradient, which he describes as "differences in stability between language components/domains (or subcomponents/subdomains), such as the difference between lexicon (less stable) and grammar (more stable)" (2000:50). The VC Model recognizes the fact that in all instances of contact-induced change, there is a source language (SL) and a recipient language (RL). The outcome of contact-induced change depends on which language is the SL and which language is the RL. This makes it possible to describe influence going in both directions. RL Agentivity involves change in the speaker's more dominant language while SL Agentivity involves change in the speaker's less dominant language. Neutralization occurs among balanced bilinguals who are not significantly more dominant in one language over another.

For example, lexicon is very easy to borrow into one's dominant language because the lexicon is a less stable domain of language. Grammar (including phonology, morphology, and syntax), on the other hand, is much more difficult to borrow because grammar is a much more stable domain. Thus, when loan words are borrowed, they tend to be modified to fit the phonological system of the RL. As a result of the stability gradient, influence from a less

---

has been used by different researchers in different ways. For example, Thomason and Kaufman (1988) also use the term *borrowing (under maintenance)*, but their usage of *borrowing* is broader than van Coetsem's.

[7] Another term used by Van Coetsem (1988; 2000) for SL Agentivity is *imposition*. This roughly, but not exactly corresponds to TK's Shift-induced Interference.

dominant language (SL) on a more dominant language (RL) is much more likely to be lexical than phonological or grammatical. This would be described as RL Agentivity.

SL agentivity would refer to contact-induced change going in the other direction for the same speaker. In other words, SL agentivity would refer to a speaker speaking the speaker's non-dominant language. The RL and the SL are reversed under SL agentivity. Thus, the influence would be from the more dominant language (SL) to a less dominant language (RL). Because grammar is a more stable domain, the influence of the SL on the RL in this case would be primarily grammatical or phonological. For example, speaking a less dominant language with a perceptible accent would be a case of SL agentivity. In this case, phonology is one of the most stable domains of language. Speakers would, thus, speak their less dominant language with influence from the grammar of their more dominant language under SL agentivity.

The VC Model also recognizes changes in linguistic dominance over the course of a speaker's lifetime. In such cases, van Coetsem describes an inverse relationship between acquisition and imposition (SL agentivity). As acquisition of a less dominant language increases, the stability factor weakens. Thus, as a speaker becomes more proficient in a language (the RL in this case), influence from the more dominant language (the SL in this case) decreases. The speaker then speaks the RL with less grammatical influence from the SL. If acquisition reaches a point at which a speaker becomes a balanced bilingual, and hence not significantly more dominant in one language than another, the differences between RL agentivity and SL agentivity become "neutralized". "Neutralization" is, thus, the third mechanism in the VC Model.

Van Coetsem formalizes the three mechanisms as follows:

"The stability factor is operational: A → $\underline{\text{B}}$ (RL agentivity) OR $\underline{\text{A}}$ → B (SL agentivity)

The stability factor is non-operational: A ←→B" (2000:42).

"A" and "B" represent the two languages spoken by a bilingual speaker. The language that is underlined is the speaker's more dominant language. RL and SL agentivity differ from neutralization in terms of whether or not the stability gradient is operational. If the stability gradient is non-operational, influence can go in either direction. Neutralization, thus, refers to cases in which speakers are highly proficient in two languages. In such cases, it is often difficult to distinguish between RL and SL agentivity. If a speaker is highly proficient in two languages, speakers have the ability to manipulate linguistic material in a greater variety of ways. Such bilingual speakers are, thus, not constrained by the stability gradient. When communities include many balanced bilinguals, the VC Model describes several possible outcomes. One example is the bilingual mixed language Media Lengua, which he says developed from Quechua phonology and syntax, but primarily Spanish lexicon. The Kupwar varieties (also discussed in Section 2.3.3), on the other hand, involve three different languages (Kannada, Marathi, and Urdu) that each retain their vocabulary but converge in phonology and syntax. Another example is Michif, which was formed with verbs coming from Cree (pronounced with Cree phonology) but nouns coming from French (pronounced with French phonology). Such languages, van Coetsem (2000) argues, are an outcome of Neutralization for individual speakers in bilingual communities.

3.2 THREE MODELS OF CHANGE COMPARED AND CONTRASTED

In Table 4, I summarize the similarities and differences across the three models of change discussed in Chapters 2 and 3: the TD Model, the TK Model, and the VC Model. The three models are similar to each other in that each one includes two basic mechanisms. The VC Model includes a third mechanism that essentially combines the other two. The underlying basis of the mechanisms proposed in each model, however, are very different from each other. This is due to the different perspectives taken in each model leading to different ways of describing the same sociolinguistic context involving heritage language speakers.

Table 4. Comparison of three models of language change

	TD	**TK**	**VC**
Purpose	Sound Change	Contact-induced Change	Contact-induced Change
Perspective	Community level	Community level	Individual level
Underlying basis of mechanisms	L1 (child) vs. L2 (adult) Acquisition	Inter-generational continuity in language use vs. language shift	Individual language dominance (proficiency)
Mechanism for L1 initiated change	Transmission	Internal Motivation	Internal Motivation
Mechanism for L2 to L1 lexical borrowing	Diffusion	Borrowing under maintenance	RL Agentivity (but L2 defined as non-dominant)
Mechanism for L1 to L2 influence	NOT CONSIDERED	Interference through Shift	SL Agentivity
Heritage language change	NOT CONSIDERED	High intensity borrowing under maintenance	SL Agentivity or Neutralization

3.2.1 One Type of Contact Mechanism vs. Two or Three

The TD Model was designed specifically for sound change from a speech community based perspective. Contact is placed in direct opposition to “regular sound change”. For this reason, the distinction between “transmission” and “diffusion” is a distinction based on speech community internal vs. speech community external change. Consequently, this model makes no distinction between different types of contact mechanisms or even different types of externally motivated change. As I discussed in Chapter 2, it is not even clear that the TD Model would even be applicable to the study of heritage languages because of underlying assumptions that do not hold for heritage language speakers and the communities to which they belong.

In the other two models, however, type of contact mechanism is the basis for distinguishing between different types of change. For example, in the TK Model the distinction between Maintenance and Shift is a distinction between two sociolinguistic processes. The former refers to cases of inter-generational transmission of a given language within a community of speakers while the latter refers to cases involving the influence of non-native speakers on a target language. Similarly, in the VC Model, the distinction between different mechanisms is based on which language is the linguistically dominant language and the direction of influence. RL agentivity involves change in the dominant language while SL agentivity involves change in a non-dominant language. Thus, unlike the TD Model, both the TK and VC Models recognize the possibility of non-native speaker influence on a target language.

Although “diffusion” (TD Model) and Shift (TK) share the common property of being based on adult language acquisition, they do not refer to the same direction of influence. “Diffusion” (TD Model) is borrowing into a language that continues to be transmitted to subsequent generations of speakers. This would be “borrowing under maintenance” in the TK Model or RL agentivity in the VC Model. “Interference through shift” or SL agentivity, however, refers to a process initiated by non-native or non-dominant speakers of a language. This is a direction of influence that the TD Model completely ignores. Both “transmission” and “diffusion” in the TD Model, thus, would be processes that occur under Maintenance (TK Model) or RL agentivity (VC Model). Interference through shift or SL agentivity are also the mechanisms that give rise to phonological influence from one language to another. The exclusion of such mechanisms under the TD Model, thus, creates no room for recognition of possible contact-induced phonological change with the exception of mergers (as discussed in Ch. 2). Instead, the TD Model treats contact-induced phonological changes only as the indirect result of lexical borrowing.

Although Labov does acknowledge some cases of interference through shift (or "substratum effects" in his terminology) such as in the development of Irish English (2008:315–316), an example also mentioned above in Section 3.1.1, he has also described "mysteries of the substrate" in the development of North American English dialects (2008:315). The first mystery is why substrate effects do not appear in many cases in which they would be expected to most likely occur. In many of the North American cases Labov (2001; 2008) has considered, sex and social class are stronger predictors of variation and change than ethnicity, which would be the variable most closely tied to substratum effects. In some cases, Labov has even observed "reverse ethnic effects" such as in the raising of /æ/ led by ethnic Italians or the raising /ɔ/ led by Jewish New Yorkers (2008:317–318). Neither of these are features found in the heritage languages of these speakers. Thus, children of immigrant parents sometimes appear to avoid variants associated with the non-native speech of their parents.

The second mystery of the substrate that (Labov 2008) discusses is why there are cases in which ethnic effects are found but the change involved cannot be tied to a specific feature in the heritage language (see also Labov 2001:247). One example is Herold's (1990; 1997) research on the low back merger in Eastern Pennsylvanian historic mining communities. In my review of this study in Chapter 2, I explained how this could arise from a smaller inventory found across many of the heritage languages spoken among the immigrants who settled in this region. Although this is consistent with Herold's (1997) explanation, Labov has expressed skepticism of this interpretation (2008:321).

It could be due to these unresolved issues that Labov does not consider interference through shift (or SL Agentivity) as part of the TD Model. Instead, Labov theorizes a close link between linguistic processes and their possible outcomes. For instance, in the TD Model, adult language acquisition is tied to lexical diffusion with irregular grammatical patterning and sometimes changes that lead to simplification such as vowel mergers. Child language acquisition, on the other hand, is tied to a greater range of possible outcomes including phonetically regular chain shifts, splits, and mergers. The linguistic outcome of Shift (or SL Agentivity), however, is also theorized as grammatically regular because of the effect of the stability gradient. What this means is that without sufficient evidence, it can be difficult to distinguish between internally motivated and externally motivated changes that arise from interference through shift (or SL Agentivity) if one focuses exclusively on data that comes from the outcome of such purported changes.

Thomason and Kaufman (1988) also discuss skepticism of explanations based on interference through shift for a different reason. In historical linguistics research, the consequences of language shift present a methodological problem. If a community undergoes complete language shift, the language as previously spoken by the community disappears. Many historical cases of shift lack documentation of what the previously spoken language was like. Although in some cases, the language previously spoken by a group continues to be spoken elsewhere, using data based on the language as it is currently spoken elsewhere would require assuming that this variety is more or less identical to the variety spoken by the pre-shifting group. Thus, without actual data on how the language was previously spoken prior to shift, it becomes methodologically difficult to build a case for interference through shift as the cause of change. Further complicating matters is the fact that contact-induced change does not always result in an exact replication of the source language feature as (Labov 2008) and (Johanson 2008) have noted. Labov also notes a similar debate in creole linguistics between explanations based on innateness accounts and explanations based on substratist accounts (2008:216).

3.2.2 The Underlying Bases of Different Models

There are also major differences in the underlying mechanisms proposed under each model. Both the TD and the TK models treat child vs. adult language acquisition as the cognitive basis for giving rise to two different sociolinguistic processes. In the TD Model, L1 acquisition is treated as synonymous with child language acquisition and L2 acquisition is treated as synonymous with adult language acquisition. L1 acquisition is also the basis of "transmission" while L2 acquisition is the basis of "diffusion". In the TK Model, on the other hand, both the L1 and L2 can be languages acquired during childhood. This recognizes cases in which the first language is not the linguistically dominant language such as in the case for many heritage language speakers. The distinction between child and adult language acquisition, however, is still recognized as the cognitive basis for what gives rise to the distinction between Maintenance and Shift. The VC Model takes a different perspective by focusing on language dominance in individual speakers rather than on relative age of acquisition for an entire community.

The different underlying bases for the different transfer mechanisms have implications for how each model would describe heritage language change. As discussed in Ch. 2, heritage language speakers who subsequently become dominant in a second language are problematic for the TD Model. Both the TK and the VC Models, however, do recognize such groups of speakers. In the TK Model, heritage languages would unambiguously fall under high intensity Maintenance. This is because speaking a heritage language involves continuing the use of a language that has historically been spoken within a population of speakers. If there is change, change can either be internally motivated or influenced by the societally dominant language in which it is spoken. Even if there is structural influence from another language, what is important in the TK Model is that a heritage language continues to be spoken and is hence maintained across multiple generations. For the VC Model, the focus on the individual means that the mechanism involved depends on the proficiency of the individual speaker rather than on which language is socially dominant within a community. Nevertheless, societal dominance "may influence linguistic dominance" (van Coetsem 2000:57). Thus, it would still be possible to discuss heritage language speakers as a group in the VC Model in cases in which the dominant language for an individual is also the societally dominant language.

Van Coetsem describes two possibilities for speakers of heritage languages in North America. He says that the most typical case involves SL agentivity on the individual level and an inter-generational shift in the dominant language leading to a change in the direction of influence. He formalizes the distinction between the two generations as follows, with A and B indicating the two languages and the underline indicating the linguistically dominant language:

"initial generation(s): L1 ($\underline{A}$) → L2 (B) = imposition by A (acquisition of B)

subsequent generation(s): L1 ($\underline{B}$) → L2 (A) = imposition by B (attrition of A)" (van Coetsem 2000:172).

As stated above, the initial generation of speakers is linguistically dominant in Language A. For this generation, the transfer mechanism involved is imposition (SL agentivity) of

Language A as speakers acquire Language B. For the subsequent generation, SL Agentivity is also involved but goes in the opposite direction. In other words, the second generation becomes dominant in Language B and imposes structural material of Language B as speakers undergo attrition of Language A. Van Coetsem notes that SL agentivity is only one possibility for individual heritage language speakers. He says, “attrition does not have to occur and is therefore not a necessary development in the second stage. Language A may be maintained and with symmetrical bilingualism neutralization may result” (van Coetsem 2000:172, FN 1). Thus, under the VC model, two possible transfer mechanisms may be involved among heritage language speakers depending on individual speaker proficiency in the heritage language.

Van Coetsem also discusses a caveat related to the duality of patterning, a concept introduced by Hockett (1958) to refer to the distinction between the phonemic system and the grammatical system of any human language. Van Coetsem says that “accent has very much a status of its own in the acquisition process, which we must recognize in order to evaluate that process adequately” (2000:177). Thus, as widely recognized by many researchers, patterns and processes affecting the phonemic system of a language can often be distinct from the patterns and processes affecting grammatical processes of a language. Van Coetsem also says,

> “Insofar as the SL speaker does not completely imitate the RL, including the latter’s articulatory habits, he has not completed the acquisition process, but insofar as he has otherwise achieved full proficiency in the RL, he has completed the acquisition process. This seemingly paradoxical statement characterizes the very nature of accent” (2000:177–178).

When discussing the stability of the pronunciation components of a language even with complete acquisition of the target language, van Coestem does not make a distinction between SL agentivity involving those who acquire a second language as adults and SL agentivity involving heritage speakers who acquire the societally dominant language as children. Whether or not the same process is actually involved is not a question that this dissertation is designed to address. What is clear based on studies discussed in the next section, however, is that even if the same process is involved, the linguistic outcomes are different. Neutralization could also provide a better account, but the exact mechanism could also vary by individual speaker. Linguistic dominance, age of acquisition, and the duality of patterning are all factors that need to be considered in order to account for heritage language speech production patterns. Although these variables may be able to address the actuation of change among individual speakers, the relationship between the actuation and propagation of change is not clear from the VC Model.

3.3 HL BILINGUAL PHONETICS AND PHONOLOGY

In this section, I discuss studies of the phonetics and phonology of heritage speakers and other types of early bilingual speakers. Many, but not all, of the studies discussed in this section are based on experimental approaches. The goal of this section is to show that any theory of contact-induced sound change must consider at least three factors (and possibly more) including the duality of patterning, age of acquisition, and linguistic dominance.

As I discussed in the previous section van Coetsem (2000) argues that the duality of patterning (Hockett 1958) explains cases of speakers who become linguistically dominant in the grammar of a second language as adults, but retain an accent in their speech. He says that this illustrates how the phonemic system of a language can be distinct from the grammatical system of a language. Mastery of the latter does not require mastery of the former. The studies I present in the following paragraphs show evidence for this duality of patterning in the speech of heritage speakers and other early bilingual speakers. The specific characteristics, however, are different from those involving adult second language speakers.

First of all, many studies of heritage language speech from a psycholinguistic perspective have shown evidence for differences between heritage speakers and monolingual speakers of the same language in morphosyntactic features. Many researchers have even defined heritage languages in terms of structural loss or incomplete acquisition.[8] For example, Polinsky defines HLs as languages "spoken by early bilinguals, simultaneous or sequential, whose home language (L1) is severely restricted because of insufficient input" (2011:para. 1). Yet, many of the same researchers who mention attrition or language loss also note the stability of the phonemic inventory of HL speakers. Polinsky & Kagan, for example, state that although HL phonetics and phonology is under-researched compared to HL morpho-syntax, "heritage speakers generally sound so native like – one could easily imagine that there would be no differences in phonological representations between the heritage language and the baseline, *although that remains to be shown* [my emphasis]" (2007:378). This difference between maintenance of phonemic contrasts and change in grammatical patterns supports the relevance of the duality of patterning.

An important clarification needs to be made about morpho-phonological processes. Where do they fit in terms of the duality of patterning? Hockett describes the duality of patterning as follows:

---

[8] See (Nagy 2015) for a critique of the "deficit" perspective even for morpho-syntactic features.

> "Any utterance in a language consists of an arrangement of the phonemes of that language; at the same time, any utterance in a language consists of an arrangement of the morphemes of that language, each morpheme being variously represented by some small arrangement of phonemes. This is what we mean by "duality": a language has a phonological system and also a grammatical system" (1958:574).

In other words, a language has a system of organizing phonemes and a separate system of organizing morphemes to create meaningful utterances including complete sentences. Hockett defines a "phonological system" as "a stock of phonemes, and the arrangements in which they occur" (1958:137). He also makes a distinction between a "phonological system" and a "morphophonemic system" and defines the latter as "the code which ties together the grammatical and the phonological systems" (ibid.). Hockett's original definition is important to clarify since different researchers interpret "phonological system" and "phonology" in different ways depending on the school of thought. For instance, unlike (Hockett 1958:137), some researchers might include morphophonemic processes such as vowel harmony, stress reduction, or devoicing of word-final morphemes as part of phonology.

Some studies do, in fact, show evidence of cross-linguistic influence of morpho-phonological processes. This suggests that these processes pattern more like the grammatical part of the duality of patterning. Lyskawa et al. (2016), for example, have shown the transfer of English de-voicing constraints to the Heritage Polish spoken in Toronto. They describe the outcome as a system that combines the de-voicing rules of Polish with the phonetically conditioned de-voicing constraints of English such that Heritage Polish has more overall devoicing than either Homeland Polish or Toronto English. (Ronquest 2013) suggests the possibility of influence of English phonological rules to HL Spanish speakers in Chicago. She found that the HL Spanish speakers produce more centralized vowels in unstressed than in stressed syllables. This could be influenced by vowel reduction in English. Thomason and Kaufman's (1988) example of the transfer of vowel harmony rules from Turkish to Asia Minor dialects of Greek is another example illustrating the possibility of cross-linguistic influence of morpho-phonological processes.

Although phonological processes and constraints are not the focus of this dissertation, these aforementioned studies do show evidence that they can be influenced by the dominant language and hence by SL Agentivity. The studies discussed in the rest of this section, however, show lack of phonemic change. Although phonetic differences have been observed, none of these studies have observed changes that lead to loss in phonemic contrasts between different vowel categories. If SL Agentivity is involved, it appears to apply only to the phonetic production of phonemes. Thus, consistent with (Hockett 1958), this supports a distinction between the phonemic system and the morphophonemic system. SL Agentivity affecting the latter does not appear to affect the former.

Yang (2014)[9], for example, addressed how two distinct vowel systems develop among HL speakers of Mandarin (who subsequently learn English) based on two different longitudinal studies. In both studies, the speakers examined included children with exposure only to Mandarin up to around the age of 3 years. These children subsequently entered preschools in the US where they encountered their first heavy exposure to English. Yang (2014) found that in the initial stages of exposure to English in the preschools, the Mandarin speaking children had an English vowel system that showed heavy influence from Mandarin. As time progressed, however, the children showed increasing separation between their Mandarin vowels and their

---

[9] I would like to thank Marjorie Chan for this reference.

English vowels. After 3 years of exposure, the children reached native monolingual English speaker targets for all English vowels. Yang (2014) notes that this is quite a rapid change in contrast to what has been observed in adult L2 phonological development. Yet, with higher proficiency in English, the children also began to show evidence of bi-directional influence such that their Mandarin vowels also changed under influence from their developing English system. Overall, however, the Mandarin system of these children maintained the same set of phonemic contrasts found among adult monolingual speakers of Mandarin even as they began to diverge phonetically from native monolingual speakers of Mandarin.

Some research has also shown evidence that whether or not the dominant language has an effect on HL phonetics depends on the type of sound involved. One pioneering study that illustrates such differential effects is Mack (1990)[10], which examined both VOT and vowel production among French-English bilingual children. Among native monolingual speakers, French voiced stops are phonetically pre-voiced while English voiced stops in word-initial position are frequently produced with short-lag VOT rather than with pre-voicing. For the bilingual speakers examined in Mack (1990), however, the voiced stops in both languages were short-lag and hence converged with more English-like pronunciations. This illustrates uni-directional influence from English to French.

The voiceless stops, on the other hand, behaved quite differently among the bilingual speakers in this study. For monolingual speakers, French voiceless stops have short-lag VOT while English voiceless stops are long-lag. For the bilingual speakers in this study, the average VOT for voiceless stops was longer for both languages. The French VOT appeared to be at a

[10] I would like to thank Pavel Trofimovich for pointing me to this reference.

compromise value between the VOT of voiceless stops among monolingual French speakers and the VOT of voiceless stops among monolingual English speakers. Interestingly, the VOT of voiceless stops in English among the bilingual speakers was even longer than that of monolingual English speakers. This appears to be a way of maximizing the difference between the VOT values for both the French and English spoken by these bilingual speakers. Thus, even though there was evidence of phonetic assimilation from English to French for one category of sounds, the overall phonemic contrast between voiceless and voiced stops was maintained for both languages. Vowels, on the other hand, appear to behave differently. The same study also examined voicing-conditioned vowel duration. Unlike the results for VOT, the results for the vowels showed evidence for bi-directional effects such that vowel production in both languages differ from that of monolingual baseline speakers of these languages.

Flege has said that the results from Mack (1990) support Hypothesis 6 of his Speech Learning Model (1995:242). According to this hypothesis,

> "the phonetic category established for L2 sounds by a bilingual may differ from a monolingual's if: 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's" (Flege 1995:239).

Maximizing the difference between the VOT values for both the French and English spoken by the bilingual speakers in Mack (1990) would, thus, be an example of "deflection". Flege also says that this process is analogically similar to historical sound change. He says, "as languages change, the raising of vowel A may precipitate the raising of B, which then causes C to rise. As the result of such push chains, the vowels A, B, and C may be produced differently, while the contrasts between them are preserved" (Flege 1995:242). What this suggests is that there may even be a connection between the cross-linguistic deflection of vowel phonemes among bilingual speakers and vowel chain shifts. Such chain shifts, of course, are primarily changes that are phonetic, but such changes can lead to vowel mergers as shown by variationist research.

In another study, Godson (2004) showed that some HL vowels may be more affected by the dominant language than others. This study examined HL Western Armenian speakers in southern California who speak English as a dominant language. Results showed evidence of cross-linguistic phonetic assimilation with English in the production of Western Armenian /i/, /ɛ/, and /a/ but not in the production of /u/ and /o/. Godson (2004) argues that this is because /i/, /ɛ/, and /a/ have phonetically similar counterparts in California English while western Armenian /u/ and /o/ lack similar counterparts. In California English, both /u/ and /o/ are fronted while these two vowels are retracted in Western Armenian. The outcome is a Western Armenian vowel space that is different from the vowel space of monolingual Western Armenian speakers. Again, it must be emphasized that these changes are phonetic. They do not lead to phonemic changes in the Western Armenian speech of heritage speakers such as vowel mergers or vowel splits.

Further supporting the uniqueness of HL phonological development are studies that have specifically compared HL vowel space with that of adult L2 speakers of the same language. Saadah (2011), for example, compared HL and adult L2 speakers of Arabic and found that the HL speakers were distinct from both adult L2 speakers and non-heritage native speakers. This study also showed evidence that the HL speakers had acquired separate phonemic systems for Arabic and English, their dominant language. The adult L2 speakers, on the other hand, showed much stronger influence from their native language, which was also English. Baker and Trofimovich (2005) also showed the same differences between early and late Korean-English bilinguals. Early bilinguals showed evidence for bi-directional influence resulting in distinct vowel spaces for their English and Korean. Those that acquired Korean as adults, however, showed only uni-directional influence from their native English to Korean. Once again, however, influence is primarily phonetic and does not lead to phonemic changes within a language.

Chang et al. (2011) have also shown that early age of acquisition of two languages might account for heritage language speech production patterns. This study compared native speakers of English who also speak Mandarin either as a HL or as an L2 acquired as an adult. While Mandarin has a phonological contrast between two high round vowels, /y/ and /u/, English has only /u/. Phonetically, however, English /u/ is characterized by an average F2 that is intermediate between the average F2 of Mandarin /y/ and Mandarin /u/. Although the advanced L2 Mandarin speakers did acquire a phonological contrast between /u/ and /y/, they still showed phonetic influence from English in terms of a relatively high F2 for /u/. The HL Mandarin speakers, on the other hand, produced Mandarin /u/ with lower F2 and thus more closely approximated the pronunciation of native monolingual speakers of Mandarin. While the L2 Mandarin speakers showed evidence for phonetic assimilation between English and Mandarin /u/, the HL speakers appeared to be more influenced by a need to maintain a phonological contrast between /u/ and /y/. Further supporting this explanation is the fact that the HL speakers produced the greatest average F2 difference between Mandarin /u/ and English /u/. For some HL speakers, the need to maintain cross-linguistic phonetic distinctions meant producing even lower F2 than native monolingual Mandarin speakers resulting in the retraction of /u/, the exact opposite of what would be predicted if these speakers were assimilating their Mandarin system with their English system. Chang et al. (2011) argue that early exposure to two phonological systems would account for HL speakers producing greater cross-linguistic and language-internal distinctions (in both languages) than L2 speakers.

Moving beyond HL contexts, Stewart (2014) has documented a very similar phenomenon in the vowel system of Pijal Media Lengua, a bilingual mixed language that historically developed from Spanish and Quichua. Traditionally, both Media Lengua and Quichua have been described as languages with only three vowels (/i/, /u/, and /a/) while Spanish has five (/i/, /u/, /e/, /o/, /a/). Stewart's (2014) acoustic analysis of Pijal Media Lengua, however, shows evidence for up to eight different vowel categories with partial overlap. Quichua derived words with /i/ are thus acoustically distinct from Spanish derived words with /i/. Likewise, the same applies to the other vowels traditionally described as shared between Spanish and Quichua. Although Pijal Media Lengua is not a HL, the conditions under which it historically developed arguably show similarities with the conditions under which HLs developed. What is particularly remarkable about the Pijal Media Lengua case is how phonetically similar vowels in Spanish and Quichua maintained relatively distinct pronunciations for multiple generations. More recent work has shown that Pijal Media Lengua speakers are also able to perceive differences between mid and high vowels contrary to what would be expected if Pijal Media Lengua had only three vowels as has been traditionally described (Stewart 2018). The study of HL vowel spaces, thus, has important implications for developing a better understanding of the historical development of languages that have arisen in multilingual contexts.

To summarize this section, I have shown evidence that the duality of patterning, age of acquisition, and linguistic dominance all need to be considered in any theory of sound change that considers the potential contribution of early bilingual speakers. The duality of patterning has either explicitly or implicitly been recognized by many researchers. For example, Polinsky and

Kagan (2007) have noted that while many studies of heritage speakers show evidence for loss in morpho-syntactic structure, the same cannot be said for studies showing maintenance of phonemic contrasts. A review of studies showing influence of phonological rules and constraints such as vowel harmony in Asia Minor dialects of Greek, vowel reduction in unstressed syllables in Heritage Spanish, and word-final devoicing in Heritage Polish suggest that rules and constraints pattern as part of a system distinct from the phonemic system. All of the other studies discussed in this section show lack of change in the phonemic system.

Both age of acquisition and linguistic dominance also need to be considered together. For early bilingual speakers, two languages are acquired at an early age. This appears to account for how it is possible that the phonemic system shows no evidence for loss of contrasts among early bilinguals even if they become linguistically dominant in their second language. This is a different outcome from what van Coetsem (2000) has noted as the typical case of an adult becoming dominant in a second language while retaining influence from the first language. On the other hand, these studies do show possible SL Agentivity in morpho-phonological processes and in phonetic production. It is not clear, however, how these low-level phonetic differences are to be interpreted. Flege (1995) suggests that cross-linguistic deflection of phonetically similar vowel categories is analogically similar to vowel chain shifts. Could low-level phonetic differences be interpreted as changes in progress?

3.4 A VARIATIONIST APPROACH TO RESEARCH PROBLEMS

In this section, I focus on the benefits of a variationist approach in addressing problems that have developed in contact linguistics and in experimentally based studies of bilingual phonetics and phonology. The problems I address are as follows:

1) The methodological difficulty of proving contact-induced change. If there is evidence for across-the-board sound change, is it really an internally-motivated change or can it be a contact-induced change due to SL Agentivity (or structural borrowing following the TK Model)?
2) The individual speaker vs. the community. If individual heritage speakers vary in terms of whether or not and how they are affected by contact-induced change (as suggested by the VC Model), how can inter-speaker variation be tied to community-level change?
3) Interpreting low-level phonetic changes. Do low-level phonetic differences among heritage speakers have implications for the development of sound change?

The first problem is the methodological problem discussed in Section 3.2. Many researchers have expressed skepticism towards accounts of contact-induced structural change. If the outcome of internally-motivated and externally-motivated change can be identical, how would researchers be able to distinguish between the two processes? In this dissertation, I follow Nagy's (2011) proposal, which focuses specifically on heritage language variation and change and involves making four sets of comparisons: 1) an inter-generational comparison, 2) a diatopic comparison, 3) a cross-linguistic comparison, and 4) a cross-community comparison. The current dissertation will focus on previously unpublished data involving an inter-generational comparison and a diatopic (two places, Hong Kong vs. Toronto) comparison. A cross-linguistic comparison (Toronto Cantonese vs. Toronto English) will be based on results shown in Hoffman and Walker (2010) and (Walker et al. 2018). A cross-community comparison would involve addressing whether or not the same features undergo change across all heritage languages spoken in Toronto. Cantonese is the first language from the HLVC Corpus for which vowel variation has been studied. Thus, a cross-community comparison of vowel variation and change would have to be part of a future project.

The inter-generational comparison will be the main focus of this dissertation. If there are inter-generational changes identified, the follow-up question would be whether or not these changes could be contact-induced changes. This follow-up question will be addressed based on two sets of comparisons. One comparison would be with the Homeland variety of Cantonese spoken in Hong Kong. The specific question to address is whether or not the inter-generational changes identified are also changes that occur in apparent time in the Homeland variety. If they are not, a case for contact-induced change would be strengthened. Finally, the second comparison would be with the English spoken by the same group of heritage speakers. Do the inter-generational changes identified in the inter-generational comparison match the phonetic production of a potential source phoneme or allophone in Toronto English? An affirmative response to this question would also strengthen a case for contact-induced change.

The second major problem addressed in the current study is the question of the individual vs. the community in contact-induced change. This tension is reflected in the differences between the TK and VC Models. It is also reflected in the tension between first and third wave approaches to variation as discussed in the previous chapter. Some researchers including Winford (2007) have argued that the VC Model is a better model because of its focus on the psycholinguistic mechanisms of change. Although this could address the actuation of change, it is not clear how this can also address the propagation of change. The way I address this problem is through multivariate analysis. For each inter-generational difference identified, I will also address what factors favor the change. The factors considered will be discussed in detail in Chapter 5. They include social factors such as sex, age, and responses to various questions that were part of an Ethnic Orientation Questionnaire. In addition, factors related to language proficiency and language dominance will also be considered. Could factors that tie most closely to contact-induced change be shown to be the factors that best account for who leads changes that can be identified as contact-induced changes? For example, could those who use Cantonese the least often be the individuals who lead in change? An affirmative response to this question would provide further support for an account based on contact-induced change.

The third problem this study will address is the interpretation of low-level phonetic differences. Phonetics has long been treated as a domain of language that is distinct from the phonemic system of a language. The studies discussed in the previous section support this distinction in showing evidence only for phonetic differences. In some cases, there appears to be cross-linguistic deflection while in other cases, there appears to be cross-linguistic assimilation as in Western Armenian (Godson 2004). One of the important findings from the work of variationist sociolinguistics is that low-level phonetic variation can have implications for long-term sound change. Can we show that this is the case for Heritage Cantonese?

This analysis in this dissertation will be framed in terms of three types of change: mergers, splits, and shifts. While the first two have implications for the phonemic system, shifts are primarily phonetic changes that do not result in loss of phonemic contrasts. I will also follow the variationist methodology of using normalized F1 and F2 data from vowel tokens to make it possible to compare the vowel production patterns of different speakers. Unlike many of the studies discussed in this chapter, the data comes from spontaneous speech recordings.

As Nagy (2015) has observed, studies of heritage speakers based on variationist sociolinguistics methods can sometimes show results that are different from studies based on controlled experimental settings. Unlike controlled tasks, spontaneous speech is more reflective of everyday conversational behavior, is more open-ended, involves communicative intent, and allows for circumlocution or changing of conversational topic (instead of forcing participants to choose a single answer related to a specific structure or form) (Nagy 2015:324–325). Furthermore, the recruitment procedures are typically different (Nagy 2015:324). The participants analyzed in this dissertation were recruited from the entire Toronto Cantonese speaking community instead of primarily through university or educational networks as is typically the case for many psycholinguistic studies. These speakers range in age from 16 to 87. For the GEN 2 group, the age range is from 19 to 44. It could be for any of these reasons, that the results of this dissertation study differ from results obtained in other studies.

For example, many studies of heritage speakers have shown maintenance of phonemic contrasts. Could it be the case that in a spontaneous speech context, the same speakers would be more likely to merge vowels? Such differences between spontaneous speech and word list contexts, of course, have widely been reported in sociolinguistic studies and have formed the basis of the design of the sociolinguistic interview (Labov 1994). If this is the case for monolingual speakers, it could also be the case for heritage speakers. Thus, although many previous studies of heritage speaker vowels show maintenance of phonemic contrasts, the use of spontaneous speech data could paint a different picture. Similarly, could it be the case that the cross-linguistic deflection described by (Flege 1995) is less likely to occur in the spontaneous speech of heritage speakers than it is in controlled contexts? If such vowel shifting is found in this study, it would be possible to address both the linguistic, social, and individual factors that favor shifting. This would in turn make it possible to address whether such variation illustrates contact-induced sound change in progress.

3.5 CHAPTER SUMMARY AND CONCLUSION

In this chapter, I reviewed two models of contact-induced change: the TK and the VC Models. The advantage of these models over the TD Model discussed in the previous chapter is their recognition of how certain sociolinguistic settings can facilitate the development of contact-induced structural change. While the TD Model focuses on sound change, the TK and the VC Models consider all domains of language including phonetics, phonology, morphology, and syntax. The TK and the VC Models differ from each other in terms of the underlying basis distinguishing between different transfer types. Under the TK Model, it is child vs. adult language acquisition as it is in the TD Model. Under the VC Model, on the other hand, it is linguistic dominance. This recognizes the possibility that one can become linguistically dominant in a second language as is the case for heritage language speakers. Hockett's duality of patterning, however, cannot be ignored as van Coetsem (2000) suggests is the case for adults who become linguistically dominant in a second language but speak the second language with a perceptible accent.

In my review of the literature on heritage language phonetics and phonology, I discussed a pattern that shows different effects on different domains of language. While many studies of heritage language morpho-syntactic features have described structural loss or incomplete acquisition, research on heritage language phonetics and phonology has generally shown maintenance of phonemic contrasts. Where there is evidence of phonetic change in vowels, the change is towards values intermediate between those of monolingual speakers of the same language and adult second language speakers of the language. This suggests vowel shifting.

I also discussed problems in the current research literature such as the methodological difficulty of proving contact-induced change, the problem of the actuation vs. the propagation of contact-induced change, and the problem of interpreting low-level phonetic variation. I explained how the variationist approach adopted in this dissertation study can address these problems. Broadly stated, the research questions of this dissertation address whether or not there are vowel shifts, mergers, or splits initiated by GEN 2 Toronto Heritage Cantonese speakers. If there are changes observed, can they be attributed to contact-induced change? If so, what are the implications of the observed findings to models of contact-induced change?

Thomason & Kaufman have said that "what is needed is research on current or recent contact situations that permit a more ambitious analysis of sociolinguistic context than we have attempted here" (1988:213). This dissertation is an attempt to address this problem. Is Toronto Heritage Cantonese similar to other heritage languages in terms of contexts in which it is spoken? Could this lead to outcomes different from what has been observed in other heritage language contact settings? In the next chapter, I provide relevant background on the social and historical context of the community under investigation.

4.0 THE SOCIO-HISTORICAL CONTEXT

As Thomason & Kaufman have said, "the history of a language is a function of the history of its speakers, and not an independent phenomenon that can be thoroughly studied without reference to the social context in which it is embedded" (1988:4). What exactly is it about the social context that is most relevant for understanding sound change in Toronto Heritage Cantonese? I begin in Section 4.1 by describing Cantonese and its relationship to other forms of Chinese. I focus on the history of the growth of Hong Kong Cantonese as a prestige standard tied to the rapid economic development of the former British colony of Hong Kong in Section 4.2. I also discuss other important languages spoken in Hong Kong including English and Mandarin and how the influence of these languages has changed over time. In Section 4.3, I discuss the history of Hong Kong Cantonese speakers in Toronto. The most important difference between Hong Kong and Toronto is the relative societal position of Cantonese vis-à-vis English. While Cantonese is the most widely spoken language in Hong Kong, it is a minority language in Toronto where it is the mother tongue of about 5% of the population[11]. This 5%, however, still represents over 187,000 speakers making it the second largest mother tongue in the Greater Toronto Area (GTA). The linguistic implication of the different social status of these two

[11] The 5% is my estimate based on the inclusion of the ambiguous term "Chinese" on the census. The actual number of people that reported "Cantonese" was 3.2% while those reporting Chinese was 3.3%. The percentage of Cantonese speakers is thus somewhere between 3.2% and 6.5%.

languages is that bilingual English-Cantonese speakers who grew up in the Homeland are more likely to be Cantonese-dominant late bilinguals while speakers who grew up in Toronto are more likely to be English-dominant early bilinguals. While there are exceptions on the individual level, all of the speakers examined will fit the criteria of being an early bilingual if raised in Toronto and a late bilingual or Cantonese monolingual if raised in the Homeland. The three speaker groups that will be examined in this dissertation are listed in Table 5.

Table 5. Speaker Groups to be Examined

Group	Place of Residence	Acquisition of Cantonese	Acquisition of English	Societal Status of English
Homeland Speakers	Hong Kong	Acquired early	Not universal (through schooling, if acquired)	Hong Kong English (with Cantonese influence, cf. Setter, Wong, and Chan 2010)
GEN 1	At least 18 years in Hong Kong; At least 20 years in Toronto	Acquired early, dominant spoken language	Not universal (varied circumstances, if acquired)	Toronto English (with Cantonese influence, cf. Hoffman and Walker 2010)
GEN 2	Toronto	Acquired early	Acquired early as L2 but dominant language	Primary language (with lack of Cantonese influence, cf. Hoffman and Walker 2010)

4.1 WHAT IS CANTONESE?

Harrison and So have said that “although rarely claimed explicitly, Hong Kong is the greatest Cantonese city that the world has ever seen” (1997:12). Similarly, Snow has said that “Hong Kong is far and away the largest and wealthiest Chinese community in the world that speaks a dialect[12] of Chinese to the almost complete exclusion of Mandarin" (2004:1). These statements highlight a major discrepancy between what is officially recognized and what has become the common experiences of people who live in one part of the world. Hong Kong's status as the prestige center of Cantonese is rarely claimed explicitly because Cantonese lacks official recognition. What is officially recognized in Hong Kong is written language, which would be English and Standard Written Chinese, which is based on a form of spoken Mandarin, a language that is very different from spoken Cantonese in terms of vocabulary, phonology, and to a certain extent morpho-syntax. Understanding the sociolinguistic situation in Hong Kong requires disentangling spoken language and written language. This is consistent with not only the descriptive linguistics tenet that spoken language and written language are two different things but is also consistent with local beliefs about written and spoken language being two different things (cf. Snow 2004).

What "Cantonese" means and the relationship between different varieties of speech that have been called "Cantonese" is important for contextualizing the historical development of Hong Kong Cantonese. According to Yue-Hashimoto (1972; 1991), the term "Cantonese" has been ambiguously used in the English-speaking world to refer to both the dialect of the city of Guangzhou (or "Canton" based on the Portuguese spelling) and a group of dialects spoken in Guangdong (which has also been confusingly transliterated as "Canton") Province. Yue-Hashimoto (1972, 1991) reserves the usage of "Cantonese" for the former and the term "Yue" for the latter. Some Cantonese linguists have referred to the Guangzhou dialect as 'Standard Cantonese' to make it more clearly distinguishable from other Yue dialects (cf. Yue-Hashimoto

[12] Snow (2004:259)_includes an endnote in which he says, "in the Chinese context, the term "dialect" is used to refer to a variety of Chinese that differs significantly from Mandarin in pronunciation, vocabulary, and to some degree in grammar" (2004:259).

1972; Bauer and Benedict 1997). Included in the geographical reach of "standard" are the dialects of Hong Kong and Macau. Yue is one of eight major sub-groupings (or "regionalects" following DeFrancis 1984) of Chinese (which belongs to the larger Sino-Tibetan Family) recognized by Chinese linguists. Sometimes these regionalects have been referred to in English as "dialects" as Snow does in the quote cited above (2004:1) but referring to them as "dialects" obscures the fact that even within these "dialects" are many sub-dialects that exist along a dialect continuum, not all of which are mutually intelligible with each other. Yet, at the same time, invoking "mutual intelligibility", as is common in many descriptions of variation within Chinese, de-emphasizes both the historic and present-day importance of multidialectalism (or multilingualism). For example, Seiyap[13] is a sub-branch of Yue that includes dialects that are not mutually intelligible with the Guangzhou dialect. Yet, many Seiyap speakers are also bi-dialectal in both the Seiyap dialects and in the Guangzhou dialect. It is because of widespread bi-dialectalism that many Seiyap speakers identify themselves as "Cantonese" speakers even if the Guangzhou dialect is not their native dialect. In fact, many speakers of the Seiyap dialect refer to their variety as a "rural" variety of Cantonese and thus consider their dialect to be a non-standard variety of Cantonese rather than a completely distinct variety (Leung 2012).

---

[13] Existing literature on the language, history, and culture of this region has adopted an extremely wide variety of spellings. I have chosen to spell it as "Seiyap", which is a spelling that reflects the Cantonese pronunciation of 四邑. This is spelled as Siyi in Mandarin Pinyin. Other spellings found in the literature include Schleiyip and Lliyip (cf. Leung 2012). These spellings reflect how Seiyap would be pronounced in different Seiyap dialects. The initial "S" in Cantonese corresponds with a lateral fricative sound, which explains either the "schl" or "ll". Some dialects have a diphthong represented as "ei" while others merge this diphthong with IPA [i]. "Seiyap" means "four counties". Many other widely used names for the dialects spoken in this part of Guangdong are based on the names of these four individual counties. The Mandarin pinyin names of these four counties are Taishan, Kaiping, Xinhui, and Enping. Most of the other names that refer to these dialects are derived from the name Taishan. This includes "Taishanese", "Toisanese, and "Hoisanese". Leung (2012) calls it "Hoisan-wa", with "wa" meaning "language" and hence the "language of Hoisan".

It is important to note that ‘Standard Cantonese’, which is a term that has been used by many Cantonese linguists (Yue-Hashimoto 1972; Yue-Hashimoto 1991; Bauer and Benedict 1997; Leung 2012), does not mean a standard based on a written language or a standard that is politically imposed as an official language. Instead, ‘Standard Cantonese’ is a ‘standard’ that developed in Guangzhou as a common form of speech independent of writing and without any political intervention. Although the socio-economic prestige of Guangzhou as an administrative capital did give the Guangzhou dialect prestige status, the Guangzhou dialect was never codified as an official written language nor was it ever politically imposed as a standard. For these reasons, ‘lingua franca’ may be a more appropriate term since ‘lingua franca’ does not necessarily entail a common language based on writing nor does it entail a politically imposed or officially recognized spoken standard. Cantonese has also been adopted as a lingua franca spoken by speakers of other subgroups of Chinese living in Guangdong and neighboring Guangxi and Fujian Provinces including speakers of Hakka and Min (Yue-Hashimoto 1991:297). Though it was originally the speech of Guangzhou, the Cantonese lingua franca has spread to Hong Kong, which, as described in the beginning of this subsection, is now recognized as the prestige center of Cantonese. This has happened in spite of the lack of de jure recognition of Cantonese.

A point of confusion among many non-Sinitic specialists is the relationship between spoken Cantonese and written Chinese. Snow (2004) has traced the history of the development of language ideologies related to writing among Cantonese speakers. Throughout much of Chinese history, China was in a diglossic situation in which the High language was classical written Chinese and the Low language was the many vernacular, and often mutually unintelligible dialects spoken in different regions (Snow 2004:29). Classical Written Chinese, was thus an elite language learned only by the educated and used as a standard written form, regardless of one's native dialect. The situation may have been similar to the use of Latin throughout the Middle Ages in Europe. Snow says that the existence of such a diglossic system continues to have an impact on the ways in which Cantonese speakers view language. This includes "(1) a tendency to accept the idea that written and spoken language can and even should be quite different from each other … (2) a tendency to believe that while it is appropriate for people within the Chinese civilization to speak different varieties of Chinese, they should all use the same written variety. … (3) a tendency to view this distinctive (and unified) written language as a symbol of Chinese civilization" (Snow 2004:29).

The Baihua ("vernacularization") Movement in the early 20th century changed the relationship between written and spoken forms of speech throughout much of China. The problem was that a variety of spoken Mandarin was chosen as the written standard. Those that also speak Mandarin were thus able to write a form of Chinese that transparently maps onto the way they speak. This form of Chinese is now recognized as Standard Written Chinese. For Hong Kong and Cantonese-speaking regions, however, the Baihua Movement had little impact. Classical written Chinese persisted much longer, but "this development, which represented increasing colloquialization of written language in most of China, appeared merely to be the substitution of one written language for another in southern China" (Snow 2004:128). Thus, neither Classical Chinese nor Standard Written Mandarin brought a written language that was closer to the vernacular of Cantonese speakers. This, however, has not stopped grassroots efforts to create a Cantonese vernacular writing system. In fact, Snow (2004) discusses tremendous growth in the use of Cantonese specific characters. These characters are not a system completely distinct from the Standard Written Chinese. Instead, they are supplemental characters used to represent words that otherwise lack suitable forms in Standard Written Chinese. Some Standard Chinese characters are also used especially if they represent cognates used in Cantonese (though they do not always have the same meaning as in Mandarin). Snow says "that the costs of learning written Cantonese are not high – due to widespread consensus on the use of the phonetic borrowing principle" (2004:184). This make written Cantonese easy to learn for anyone who already knows Standard Written Chinese. In spite of the increasing popularization of written Cantonese in advertisements, mass media, social media, etc, written Cantonese continues to lack official recognition. Some Hong Kongers continue to have negative attitudes about their usage. Snow (2004:197) discusses one example of a teacher interviewed for a newspaper who became irate after she saw that the newspaper quoted her verbatim using vernacular Cantonese characters instead of in Standard Written Chinese. This anecdote would seem unusual in many American and European contexts where the favored expectation might more likely to be value faithful quotation of what the interviewee says. In this anecdote, however, the teacher is more concerned with the published form of her ideas than the verbatim transcription of her speech using characters that are socially stigmatized as part of working class culture. Thus, a strong belief persists that written language and spoken language are two different things.

For the linguist, what this means is that Cantonese may be a much more ideal language to study than English because of the strict ideological separation between written and spoken language. The writing system, however, still has an influence on spoken Cantonese, but mostly in terms of vocabulary, which end up being pronounced in Cantonese phonology. In Section 4.4, I further discuss the ways in which the writing system can influence Cantonese.

4.2 CANTONESE IN THE HOMELAND

The story of how the Hong Kong variety of Cantonese became the prestige variety is a story that is tied to the rapid development of Hong Kong as a major global economic powerhouse during the second half of the 20th century when many migrants from Guangdong Province sought refuge in what was at that time a British colony. Also important to this story is language and dialect shift particularly among speakers of other varieties of Chinese, who also migrated to Hong Kong. This led to an increase in the number of speakers of Hong Kong Cantonese and further contributed to the overall dominance of Cantonese as the de facto, but never explicitly de jure, official spoken language of Hong Kong. As a former British colony, English has had a major influence on the development of Hong Kong Cantonese. Now that Hong Kong is once again part of China, Mandarin has played an increasingly important role. Nevertheless, Cantonese is still the dominant spoken language and has consistently remained the mother tongue for about 90% of Hong Kong's population. In this section, I tell the story of how all of this happened.

4.2.1 The Emergence of Hong Kong Cantonese

Though not always emphasized, the history of Cantonese is one in which a history of contact between different groups of speakers of different languages and dialects at different time periods is important for accounting for changes that have occurred during different periods of its history. For example, ethnic Chinese migrated southward from northern China to what is now Guangdong Province during several different waves. Major waves occurred during the Qin Dynasty (221-206 BC), at the end of the Han Dynasty (c. 200 AD), during the Tang Dynasty

(618-907 AD), and at the end of the Song Dynasty (c. 1200 AD) (Yue-Hashimoto 1991:295). These migrations brought ethnic Chinese in contact with various indigenous groups including the Tai, the Miao, and the Yao. Some features found in Cantonese that are not found in Northern varieties of Chinese such as Mandarin have been attributed to substratum influence from speakers of these languages. Some examples include vocabulary, a tense/lax vowel system, and a modified-modifier word order for a set of compounds (Yue-Hashimoto 1991:312–313).

Even in more recent times, contact has continued to play an important role in shaping the development of Cantonese. In the past century, the three most important groups of speakers that have had an influence on Cantonese are speakers of other southern varieties of Chinese, speakers of varieties of English, and speakers of varieties of Mandarin. This is the case for Cantonese both in the homeland, defined as Guangdong Province and immediately adjacent areas including Hong Kong and Macau, and in many diasporic communities including Toronto, Canada. The focus of this section is on describing the socio-historical circumstances that have brought speakers of Cantonese in contact with groups of speakers of languages from these three groups and on the relative importance of each group of languages.

The modern history of Hong Kong, which was formerly part of Bao'an County in Guangdong, begins in the middle of the 19th century when it became a British colony through a series of treaties signed between the British government and Qing Dynasty China. First, was the Treaty of Nanking in 1843. At this time, the territory of Hong Kong consisted of only an island with small farming and fishing villages and a population between 5,000-7,000 (Carroll 2007:19). The colony expanded in 1860 to include what is now known as the Kowloon Peninsula and then again in 1898 to include what are now known as the New Territories. The 1898 expansion resulted in a ten-fold increase in the size of the territory pushing the northern end of the colony to the Shenzhen River, which now marks the boundary between Hong Kong and what is referred to as "Mainland China" (Carroll 2007:67). The treaty also stipulated that Britain would return the New Territories back to China after the end of a 99-year lease (Tsang 2007:39). Later negotiations resulted in the inclusion of Hong Kong Island and Kowloon as also part of territory to be returned to China. During these 99 years, the entire British colony underwent a dramatic transformation from a small colony of about 300,000 people to an economic powerhouse of over 6 million people making it one of the densest territories in the world. From 1901 to 1941, the population grew more than five-fold from 301,000 to 1,639,000 (Tsang 2007:109).

According to Tsang (2007), most of those who migrated to Hong Kong during this time period migrated for employment opportunities with the intent of returning to their ancestral homes. Most of these people came from Guangzhou or elsewhere in Guangdong. According to Zhang (2009), Guangzhou Cantonese gradually emerged as the lingua franca used between different groups of Chinese in the British colony because those from Guangzhou were the largest group of migrants and because of the relatively high socio-economic status of these migrants. Other groups of Chinese in Hong Kong included the Tanka, the Hakka, and the Hoklo. The Hakka and Hoklo spoke very different varieties of Chinese that belong to different regionalects.

The oldest GEN 1 and Homeland speakers grew up around the middle of the 99-year lease period. This also happened to be one of the most turbulent periods of Hong Kong history. It was a period characterized by a series of wars and invasions and general political turmoil in Mainland China. Many GEN 1 speakers describe this period as a period of frequent back and forth movement between the colony and Guangdong. The first major conflict of this period began in 1937 when Japan invaded China. By 1938, Japan had captured Guangzhou. To escape the Japanese invaders many people from Guangzhou as well as elsewhere in Guangdong

Province, migrated across the Shenzhen River to seek refuge in British-controlled Hong Kong. From 1937 to 1941, the population of Hong Kong increased by 63% or by about 600,000 people (Tsang 2007:115). On December 8, 1941, Japan invaded Hong Kong as part of the same Pacific military campaign that included the bombing of Pearl Harbor. Hong Kong then fell under Japanese control until the end of World War II. Japanese authorities forced many of the most recent migrants to return to their ancestral villages in Mainland China. As a result, Hong Kong's population dropped from 1.5 million to 600,000 during the Japanese occupation (Carroll 2007:123). Hong Kong quickly recovered this population loss after World War II when many of those who were forced out of Hong Kong moved back.

Although World War II officially ended in 1945, Mainland China remained politically unstable with a civil war between the Kuomintang ("Nationalist Party") and the Communist Party. Following the Communist Party victory in 1949, the People's Republic of China (PRC) was established as the government of Mainland China. Many of those who opposed the government sought refuge in Hong Kong as part of a mass migration that would continue for several decades. By 1955, Hong Kong's population reached 2.5 million surpassing the pre-World War II peak (Carroll 2007:140). As political and economic turmoil worsened in Mainland China, migrants continued crossing the border from Guangdong to Hong Kong. Some years had significantly more migration than other years such as the years during the Great Leap Forward (1958-1961), a societal transformation campaign that resulted in massive famines throughout the country. This brought even more migrants crossing the border from Guangdong to Hong Kong both legally and illegally (Carroll 2007:149). By 1960, Hong Kong had a population of 3 million.

Hong Kong's tumultuous war period and subsequent population growth had major sociolinguistic ramifications. As Labov has said, "World War II has been a watershed of linguistic behavior in many countries and for many linguistic variables" (Labov 2001:227). This appears to apply to Hong Kong much as it does in many other places around the world. For Hong Kong, however, the war period lasted beyond the official dates of World War II (1939-1945) because the territory was also heavily affected by the Japanese invasion that occurred before and the Communist Revolution in China, which followed shortly afterwards. Those that were born in Hong Kong after the Communist Revolution in 1949 developed a new sense of local Hong Kong identity that was different from those who were born earlier. This includes many of the GEN 1 and Homeland speakers that will be examined in this study. The sociolinguistic consequence of the new sense of Hong Kong identity has been the strengthening of Cantonese and the gradual emergence of a Hong Kong variety of Cantonese distinct from the Guangzhou variety.

Table 6. Hong Kong Population and Housing Census (1971), Adapted from Bauer (1982:19)

	Language Used at Home						
Place of Origin	**Cantonese**	**Seiyap**	**Hakka / Hoklo**	**Other Chinese**	**English**	**Other / Mute**	**Total**
Hong Kong	158,790	153	25,296	586	392	482	185,699
Guangzhou, Macau	1,983,372	2,571	74,203	7,951	1,200	2,786	2,072,083
Seiyap	632,174	42,346	4,933	4,242	244	835	684,774
Chaozhou	262,683	1,006	113,148	13,800	143	674	391,454
Other Guangdong	232,215	461	14,070	2,916	192	361	250,215
Other China	187,184	471	36,248	59,433	655	802	284,793
Other / Unknown	12,817	45	681	587	38,293	15,189	67,612
Total	3,469,235	47,053	268,579	89,515	41,119	21,129	3,936,630

Before the Communist Revolution, there was frequent back and forth movement between Hong Kong and Guangdong. This created “a sojourner mentality and … the non-development of a sense of local identity” (Tsang 2007:181) among Hong Kong residents. Most intended to stay primarily for work and were, in fact, permitted to travel back and forth freely. All of this changed after the establishment of the PRC. After the 1950s, most of those who migrated to Hong Kong decided to stay there permanently. This was partly due to new travel restrictions imposed by the Communist government.

Even with mass migration to Hong Kong during the post-war period, most migrants came from similar geographical and linguistic backgrounds. Table 6 presents data from the 1971 Hong Kong Census as reproduced by Bauer (1982:19). As shown in the census, a relatively small percentage of Hong Kong residents traced their ancestral origins to Hong Kong (185,699 out of 3,936,630 or about 4.7%). Most of these residents also used Cantonese as their primary language (158,790 or 86%). The most common place of ancestral origin was either Guangzhou or Macau (2,072,083 or 53%). Guangzhou, of course, was the historic center of Cantonese. Not surprisingly, an overwhelming majority of those who traced their origins to Guangzhou or Macau reported using Cantonese as their primary home language (1,983,372 or 96%). What is more remarkable is evidence of language or dialect shift to Cantonese. The second most common place of ancestral origin reported was Seiyap. Yet, only 42,346 (or 6.2%) of those who traced their origins to Seiyap reported using one of the local dialects of Seiyap at home. An overwhelming majority of those from Seiyap have shifted to Cantonese. This is further supported by Tsang (1984), who showed how the prestige of Cantonese motivated dialect shift among children of Seiyap-speaking parents in both Hong Kong and in North American Chinese communities during this era. The census also highlights the overall strength of Cantonese.

Regardless of place of origin, 3,469,235 (or 88%) of Hong Kong residents reported speaking Cantonese as their primary home language in 1971.

The year 1971 also represents a point in time when a new generation born and raised in Hong Kong came of age. This was a generation completely unfamiliar with life "behind the Bamboo Curtain" in Mainland China and one that benefitted from educational and economic opportunities in a politically stable environment. The increasing wealth of the colony during this time period also meant increased leisure time. This was a period when the arts, TV, radio, the movies, and music began to flourish. Tsang (2007:193) has said that the popular culture that emerged was one embraced by all social classes and age groups. More importantly, the new popular culture that emerged was one based on Cantonese. Snow (2004:139–140) has described how TV led the way. In the 1960s, less than 10% of the population had access to a TV. This percentage increased dramatically to 85% by 1973. By the 1980s, 97 percent of Hong Kongers watched TV for at least 15 minutes a day, with 93% watching TVB, a pioneering TV station that popularized local programming broadcasted in Cantonese. TVB programming is now available in over 40 countries worldwide reaching over 300 million households[14]. This is quite a remarkable figure given that Hong Kong's population is only 7 million. The increasing popularity of Cantonese TV paved the way for the development of Hong Kong cinema. For example, one of the best known celebrities from this era was Bruce Lee, who popularized the kung fu genre and brought it international recognition (Carroll 2007:160). Mass media, thus, had a profound effect on solidifying a distinct Hong Kong identity based on the Cantonese spoken by the local population.

[14] Statistics from the TVB International web site, consulted July 16, 2016: http://b.tvb.com/tvbi/

The 1980s and 1990s marked another major period of transition in Hong Kong. The Sino-British Joint Declaration was signed in 1984 and ratified the following year. According to this declaration, Britain agreed to return its entire colony including Hong Kong Island and Kowloon, which were initially ceded to Britain in perpetuity, as well as the New Territories to the PRC government in 1997. The PRC agreed that Hong Kong would be placed under a "one country, two systems" policy meaning that Hong Kong would be reintegrated as a part of China but would not be subject to the socialist system in place in Mainland China for a period of 50 years. Instead, Hong Kong would maintain many of the same systems that had already been in place under British colonial rule. In spite of the agreement that much would remain the same, many Hong Kongers became skeptical about the PRC government's promises of autonomy. Skepticism increased after the Tiananmen Square Massacre in 1989, which led to the killing of pro-democracy demonstrators in Beijing. Under a climate of fear and uncertainty about what would happen after 1997, many Hong Kongers, especially those from elite backgrounds, decided to immigrate to other countries in the 1980s and 1990s. One of the top countries during this time period was Canada. Within Canada, Toronto was one of the top destinations. It was under these conditions that many of the GEN 1 speakers examined in this project immigrated to Canada. All of the Toronto speakers examined immigrated before 1990.

For the Homeland speakers that remained in Hong Kong from the 1990s through the present, the most important sociolinguistic change has been the increasing role of Mandarin. In spite of the increasing importance of Mandarin, however, Cantonese remains as the dominant language of the region. According to the most recent survey (conducted in 2015), 88.1% of Hong Kongers speak Cantonese as a mother tongue (see Table 9). This figure is virtually unchanged from the percentage reported in 1971 (see Table 6). Furthermore, Table 10 shows that only 2.0%

of Hong Kong residents between the ages of 6 and 65 reported no knowledge of Cantonese in contrast to 12% who reported no knowledge of Mandarin and 13.4% who reported no knowledge of English. The strength of Cantonese in Hong Kong has been so strong that even some non-ethnic Chinese groups have also adopted Cantonese. For example, Pannu (1998) surveyed adolescents of Indian (South Asian) descent who are also speakers of both English and Punjabi and found that Cantonese was the most common language they used outside of home even with other Punjabi heritage language speaking peers. These adolescents also code-mix Cantonese with Punjabi. The strength of Hong Kong Cantonese is also evident in its influence on other varieties of Cantonese. A few widely cited examples include the spread of the words "baa1 si2" ('bus') and "dik1 si2" ('taxi'). These are both transportation terms borrowed from English in Hong Kong. They have now become commonly used in Guangzhou Cantonese as well, which otherwise has had very little contact with English and instead more contact with Mandarin (Snow 2004:265).

4.2.2 The Increasing Influence of English and Mandarin in Hong Kong

Though Cantonese has clearly become the dominant language across many domains in Hong Kong since the days of early settlement by traders from Guangzhou and remains so to this day, it is important to discuss how two other languages may have had an impact on life and language use in Hong Kong: English and Mandarin (both spoken and written).

English has been present in Hong Kong since the British colonial era, when it was the language of government and commerce. Although English has been an official language longer than Chinese, it was not until the 1990s that more than half of Hong Kong's population was able to converse in English (See Table 7). For most ethnic Chinese, schooling was the primary means of exposure to English, but mandatory schooling was not introduced until 1971 at the primary level and 1978 at the secondary level (Carroll 2007:159–160). Consequently, the spread of knowledge of English as well as code-mixing behavior has been class stratified. As Cheung has said about the situation in the 1980s, “while English in Hong Kong divides people into those who know the language (the middle class) and those who do not (the working class), Cantonese unites the general public, and mixed code unites the middle class” (1985:198).

Knowledge and proficiency of English continued to increase from the 1990s to the present. Joseph (1997) describes a paradox in the status of English in Hong Kong in the 1990s. On the one hand, many educators observed declining standards in the use of English. Yet on the other hand, there are many statistics such as the ones shown in Table 7 and Table 8 that show an increasing number of speakers who can speak English as well as an increase in self-reported proficiency in English. Joseph (1997) argues that this paradox can be reconciled by observing that more speakers of English does not necessarily mean that all of these new speakers become proficient in a standard variety of English. Instead, they develop a new variety of English that shows heavy substratum influence from Cantonese. More recent work has recognized Hong Kong English as an emergent post-colonial variety of English (Sung 2015; Setter, Wong, and Chan 2010). By the early 2000s, knowledge of English became common enough for virtually all young Hong Kong Cantonese speakers to code-mix at least some of the time.

Chen (2005; 2008) shows that socially distinct code-mixing styles have even emerged and that speakers can even perceive different styles as indexing different social identities. The Hong Kong Cantonese that has developed has become one in which English insertion into Cantonese is the norm. Yet, at the same time, English usage cannot be overdone. Otherwise, a speaker is perceived to be "pretentious" and part of an elite class of Hong Kongers who were able to return to Hong Kong after receiving some of their education in Anglophone countries abroad. Speaking a pure form of Cantonese devoid of English influence is also seen as unusual or even indexing lack of education such as would be the case for many recent migrants from Guangdong Province. This is finely illustrated in the title of a documentary film about code-switching in Hong Kong called "Present 一個 [jat1 go3] Project" (Chen and Carper 2005). When Hong Kongers were asked in front of a camera to provide a Cantonese translation of the English sentence "I need to present a project", all of them struggled to translate the words "present" and "project" in a way that sounded like natural Hong Kong Cantonese. This illustrates the pervasiveness of English in the everyday Cantonese spoken in Hong Kong.

The second major language in Hong Kong is Mandarin. It was not until the Official Languages Ordinance of 1974 that "Chinese" was given legal status as a second official language. The ordinance, however, did not explicitly mention a spoken variety of Chinese. Instead, written Chinese became implicitly recognized as Modern Standard Chinese, which is based on a spoken form of Mandarin (Cheung 1985:191). The de facto spoken variety of Chinese was implicitly recognized as Cantonese by virtue of the fact that most of Hong Kong's population speaks Cantonese as a mother tongue. Without explicit recognition, Cantonese became the dominant spoken language used in legal settings. For example, one study of speeches in the Hong Kong Legislative Council found that 99.45% of speeches in 2012 were in Cantonese in contrast to 100% in English in 1972 (Evans 2014).

Pierson (1998:95) notes that the ambiguity in "Chinese" may have been intentional for political reasons. In other words, "Chinese" is recognized as the official language of Mainland China and is also already the dominant language of Hong Kong although the reality is that

"Chinese" means different things from the perspective of Hong Kongers compared to the perspective of Beijing. In Hong Kong, "Chinese" means spoken Cantonese and written Mandarin while in Beijing it means both spoken and written Mandarin. Thus, not specifying whether "Chinese" in a legal document refers to a spoken language would have been a way of crafting legislation amenable to all parties. This plays around with the ambiguity of what "Chinese" (中文) means in both English and in Chinese.

Since reintegrating with the rest of China in 1997, there have been both increased efforts to teach Mandarin in the school system as well as more Mandarin-speaking migrants from other parts of China living in Hong Kong. This contrasts with the years immediately following the Communist Revolution. Although there was a small population of Mandarin speakers and many popular films and music produced in Mandarin in the 1950s, Cantonese increased its dominance from the 1960s to the 1980s (Snow 2004:179). During these years, schooling would have been the primary means of exposure to spoken Mandarin as was the case for English. In the 1950s, however, Mandarin teachers were lacking to the point that Mandarin was eventually removed from the curriculum in primary schools (Leung and Wong 1997:35). Mandarin was not re-introduced into the school curriculum until the 1980s and even so it initially began as an extra-curricular subject in primary schools (ibid.). Since then, the percentage of Hong Kongers proficient in Mandarin has increased. Yet, throughout much of this time period Mandarin has been a third language in Hong Kong with English being second. It has only been recently that the number of speakers proficient in Mandarin has begun to surpass the number of speakers of English. The most recent survey shows that 13.4% of Hong Kong residents between the ages of 6 and 65 have no knowledge of English compared to 12% who have no knowledge of Mandarin (see Table 10). Still, this is a relatively small difference.

Table 7. Language use survey results (1983 and 1993) reported in Joseph (1997:62)

Language	Can Speak (1983 Survey)	Can Speak (1993)	Can Understand (1993)
Cantonese	98.5%	91.9%	91.5%
English	43.3%	65.8%	68.6%
Mandarin	31.9%	55.6%	61.9%

Table 8. English proficiency survey results (1983 vs. 1993) reported in Joseph (1997:63)

"How well do you know English?"	1983	1993
'Quite well'/ 'Well' / 'Very well	5.1%	33.7%
'Not at all' / 'Only a few sentences' / 'A little'	92.8%	66.3%

Table 9. 2015 Survey of Hong Kong Residents by Age Group and Mother Tongue[15]

Mother Tongue	6-14	15-24	25-34	35-44	45-54	55-65	Overall
Cantonese	89.3%	91.8%	85.8%	83.0%	89.7	90.0%	88.1%
Mandarin	2.7%	3.0%	5.6%	6.5%	3.1%	2.1%	3.9%
Other Chinese	< 1%	1.7%	2.8%	4.2%	4.5%	5.9%	3.7%
English	3.8%	0.6%	1.8%	2.5%	0.8%	0.5%	1.4%
Other Asian Languages (including Filipino and Indonesian)	3.3%	2.3%	3.0%	3.2%	1.4%	1.1%	2.5%
Other	< 1%	< 1%	1%	< 1%	< 1%	< 1%	< 1%

Table 10. 2015 Survey of Perceived Language Competence for Persons Aged 6-65

	Very Good	Good	Average	Not So Good	No Knowledge
Cantonese	54.9%	31.7%	9.1%	2.4%	2.0%
Mandarin	4.8%	19.9%	41.1%	22.1%	12.0%
English	4.6%	18.6%	41.8%	21.6%	13.4%
Written Chinese	30.9%	42.1%	20.9%	3.1%	2.9%
Written English	4.6%	18.7%	41.6%	20.8%	14.3%

[15] From the Hong Kong Thematic Household Survey Report No. 59 http://www.statistics.gov.hk/pub/B11302592016XXXXB0100.pdf

4.2.3 English or Mandarin phonetic/phonological influence?

The increasing importance of both English and Mandarin in Hong Kong raises the question of whether or not these languages can have phonetic or phonological influence on Cantonese in Hong Kong. According to the models of contact-induced change presented in the previous chapter, such influence seems unlikely. First of all, Cantonese remains the mother tongue and dominant language for almost 90% of Hong Kong residents. According to the VC Model, Hong Kong Cantonese speakers would use RL Agentivity to borrow words from their non-dominant languages (English or Mandarin). SL Agentivity resulting in structural influence would apply only on the English or Mandarin spoken by Hong Kong Cantonese speakers. When words are borrowed into Cantonese from these languages, these borrowings do not lead to phonetic or phonological change.

One example to illustrate the borrowing process is the word for “fruit”. In Cantonese, the word for “fruit” is pronounced [saŋ.gʷɔ] and is written orthographically as 生果. These two characters are also used in Mandarin. A Mandarin speaker would read these characters as [ʂəŋ.gʷuo]. Since it is not the normal word for “fruit” in Mandarin, a Mandarin speaker may not necessarily be able to interpret these two characters as referring to a generic term for ‘fruit’. It would still makes sense to a Mandarin speaker, but might be interpreted differently. For instance, a Mandarin speaker could take a literal interpretation and read the word as “raw fruit”. [saŋ] in Cantonese does also mean “raw” but [saŋ.gʷɔ] is a recognized Cantonese compound that is generically used to refer to all fruit. The word for “fruit” in Mandarin is actually [ʂuei.gʷuo] or 水果. These two characters can be pronounced in Cantonese as [sɵɥ.gʷɔ]. Again, a literal interpretation is possible for the Cantonese speaker, who might interpret the Mandarin form as "water fruit". The use of [sɵɥ.gʷɔ] instead of [saŋ.gʷɔ] for Cantonese speakers illustrates a lexical borrowing from Mandarin to Cantonese. What is important to note about this borrowing is that it is mediated by the writing system. It is not a borrowing based directly on the pronunciation in the source language. Since this word was borrowed in this manner, no phonetic or phonological material from the source language was borrowed along with it. What this suggests is that even if Cantonese has borrowed a lot of vocabulary from Mandarin, the borrowing is strictly lexical. The borrowed form has been converted to Cantonese pronunciation patterns based on how it is written in Standard Written Chinese.

The writing system can also be involved in mediating the pronunciation of English loan words, including proper nouns and place names. For example, "Los Angeles", has been transliterated in at least two different ways in Written Chinese. The first is 羅省. This would be pronounced as [lɔ.saŋ] in Cantonese but as [luo.ʂəŋ] in Standard Mandarin. The second way of transliterating "Los Angeles" is as 洛杉磯, which would be pronounced as [lɔk.t͡sʰam.gej] in Cantonese but as [luo.ʂan.t͡ʃi] in Mandarin. The first transliteration happens to be based on the Cantonese pronunciation (Dong and Hom 1980:6). The second one is now recognized as the Standard Written Chinese transliteration and is also replacing the first one as the recognized Cantonese name for "Los Angeles" according to CantoDict[16]. It is clear that the Mandarin pronunciation of these three characters is much closer to the English pronunciation of "Los Angeles" than the Cantonese pronunciation. The Cantonese pronunciation includes sounds such as [k] and [m] that seem puzzling if one did not know about how the Chinese Writing System mediates pronunciation. Transliterations based on one dialect do not always take into

[16] Cantodict is an online crowdsource dictionary: http://www.cantonese.sheik.co.uk/scripts/wordsearch.php

consideration how they are pronounced in another dialect. With the increasing influence of Mandarin, [lɔk.t͡sʰam.gej] has become increasingly common among Cantonese speakers. In this case, this can be described as a borrowing from Mandarin mediated by a Standard Written Chinese transliteration of an English name that originally came from Spanish. It is not, however, a phonetic borrowing form Mandarin. In fact, no new pronunciation patterns are introduced into Cantonese when such Mandarin borrowings are introduced. When Cantonese speakers learn Standard Written Chinese, they learn the same writing system as Mandarin speakers, but they learn Cantonese specific pronunciations for each individual character. Many of these characters correspond to words inherited from Middle Chinese (an ancestor language shared by both Cantonese and Mandarin) that are pronounced one way in Modern Cantonese but in a different way in Modern Mandarin. Through this process of borrowing from Mandarin via the writing system, a case for Mandarin pronunciation influence on Cantonese seems weak.

English loan words that have entered Cantonese directly rather than mediated by the writing system also show lack of evidence of phonetic or phonological change in Cantonese. Bauer (1985) shows that when English loan words enter Cantonese, they are typically adapted to fit into the existing phonotactics of Cantonese. In some cases, English loan words introduce phonotactic sequences that otherwise do not exist in Cantonese vocabulary due to accidental inventory gaps. In some, cases, they can even violate what are believed to be phonotactic constraints in the language. One example that Bauer (1985) discusses is a phonotactic constraint against have both a labial onset and a labial coda in the same syllable. Such words are completely absent in the native Cantonese lexicon. The English word for “pump”, however, has been borrowed into Cantonese as [pɐm]. This is a word that violates the phonotactic constraint against having a labial onset and a labial coda in the same syllable. Since loan words such as [pɐm] can sometimes create phonotactic sequences that are otherwise unattested in Cantonese, Bauer (1985) describes the outcome as one that leads to an expansion in the Cantonese syllabary. The extent to which this leads to long-term change in Hong Kong Cantonese phonology remains an open question.

4.3 CANTONESE IN TORONTO

Bacon-Shone and Bolton have said that “the boundaries of the Hong Kong speech community now extend overseas to North America, Australia, and the UK; and English is obviously one of the languages linking this extended community together” (1998:84). As I will show in this sub-section, the social history of Cantonese in Toronto is in many ways a continuation of the social history of Cantonese in Hong Kong. The prestige of Cantonese in Hong Kong carried over to Toronto. What is different between the two cities is in the role that English plays. English has become the dominant language for the GEN 2 speakers interviewed for the current project, as evidenced from responses to the EOQ survey. This raises the possibility of structural influence from English to Cantonese, which we are less likely to expect in Hong Kong.

Extending the notion of “heritage prestige” introduced by Kiesling and Wisnosky (2003), I argue that Cantonese has “heritage prestige” in the Toronto context and that this heritage prestige is strengthened by the socioeconomic status of many of the Hong Kongers who immigrated to Canada beginning in the 1960s. The heritage prestige of Cantonese coexists with the societal prestige of English. This makes a description based simply on societal linguistic dominance an oversimplified one since such a model fails to recognize the possibility of two distinct norms. Furthermore, language maintenance does not necessarily mean lack of sound change. The societal status of the two languages creates the possibility of cross-linguistic influence.

The social history of Cantonese speakers in Toronto that is most relevant to this dissertation spans from the 1960s through the 2000s. As I discuss in the following paragraphs, these years were characterized by a growth in immigration from Hong Kong. I will show that the relatively high socioeconomic status of many of these immigrants has contributed to language maintenance and greater opportunities for exposure to Cantonese than is typical in many other cases of heritage languages examined by researchers[17]. Prior to the 1960s, the most widely spoken varieties of Chinese in many North American Chinese communities including Toronto were Seiyap dialects (Thompson 1989:50). This would quickly change after changes in immigration laws.

From 1966 to 1981, the Chinese population in Toronto increased from 8,000 to 80,000 making it the fasting growing Chinese community of any North American city at the time (Thompson 1989:5). What was formerly one of the smallest urban Chinese communities in North America has now become one of the largest. From 1968 to 1979, 85% of Chinese immigrants to Canada came from Hong Kong (Thompson 1989:152). The newcomers from Hong Kong quickly outnumbered the original community consisting primarily of those tracing their origins to Seiyap. Along with increasing numbers of Hong Kong immigrants came increasing numbers of

[17] For instance, Polinsky defines heritage languages as languages "spoken by early bilinguals, simultaneous or sequential, whose home language (L1) is severely restricted because of insufficient input" (Polinsky 2011:para. 1). This definition does not seem appropriate for the social context of Toronto Heritage Cantonese, but it does reflect the typical contexts studied by many researchers who focus on heritage languages. Although exposure to Cantonese for GEN 2 speakers in Toronto is still less than what would be the case if they were in Hong Kong, "severely restricted" appears to be too exaggerated of a statement to make for GEN 2 Toronto Heritage Cantonese speakers.

Cantonese speakers in Toronto. Other sources of growth in the number of Cantonese speakers included immigrants from Guangdong Province and ethnic Chinese from Vietnam. Dialect shift within Seiyap speaking families also contributed to the growth of Cantonese much as it did in Hong Kong. In many North American cities, those that had Seiyap speaking parents shifted to English while those that did continue to speak a variety of Chinese typically learned to speak Cantonese rather than Seiyap dialects (Tsang 1984; Leung 2012). This shift has been motivated by the prestige tied to the higher socio-economic status of Hong Kong immigrants (ibid.).

What made the large influx of Hong Kong immigrants possible was changes in Canadian immigration law. The first major change occurred in 1962 when "country of origin" was removed as a selection criterion (Chan 2011:125). This ended discriminatory laws that severely limited the number of immigrants from East Asia and many other parts of the world. Since then, many subsequent changes to immigration law have generally given those from Hong Kong an advantage especially for those from elite backgrounds. For example, the 1967 immigration law created a points system that assessed each person applying for immigration based on training, education, occupational skills, knowledge of English and French, employment opportunities, pre-arranged employment, and personal qualities (Chan 2011:125). At this time, Hong Kong was becoming an emerging global economic powerhouse. The increasing wealth of the colony along with an increasingly educated workforce paved the way for an increasing number of Hong Kongers who were able to benefit from the points system.

The Immigration Act of 1976 further expanded the number of immigrants allowed by establishing four classes of immigrants (refugees, families for family reunification, assisted relatives, and independent immigrants) as well as a quota system for each of these classes that could be adjusted annually based on the current global political climate (Chan 2011:125). The “family” and “assisted relatives” classes made possible immigration from a more diverse group of Cantonese speakers. While the “family” class made it possible to reunite nuclear families in Canada, the assisted relative class made it possible for even extended family members to immigrate including brothers, sisters, aunts, uncles, nieces, nephews, and grandparents. These two classes were important in creating a community that was more than simply highly educated and wealthy immigrants. In fact, there were some years, such as during the 1980s, in which more than half of all Hong Kong immigrants entered Canada through family connections. Although this percentage declines to about 40% by the 1990s, this was still a relatively large percentage of immigrants that did not necessarily come from an educated or wealthy background. For example, many older women who came under the family reunification and assisted relative classes lacked English language skills and were able to secure only low-wage jobs in Chinatown restaurants (Chan 2011:136). These immigrants were important in providing the labor needed to run many businesses started by wealthier immigrants. The fact that not everyone came from the same background or had the same professional skills contributed to the diversification of the Cantonese-speaking community in Toronto. In addition, some of the newcomers were Cantonese-speakers from Guangdong or Vietnam who had relatives in Hong Kong. Thus, even those who did not come from Hong Kong were able to contribute to the growing size of the Cantonese-speaking community in Toronto. Cantonese speakers in Toronto are anything but a homogeneous group. Immigration policy benefitting family reunification changed Toronto’s Chinese community from one that consisted largely of male bachelors from the Seiyap region to one that was much more diverse overall than ever before. By 1971, 83 percent of the Chinese population in Canada lived in a nuclear family. The Toronto Chinese community that emerged included those from a wide range of professions including "artists, chefs, filmmakers, intellectuals, journalists, merchants, physicians, social workers, and writers" (Chan 2011:121).

While it would be over-simplistic to characterize the post-1960s Toronto Cantonese speaking community as a middle and upper class community, it would be more accurate to say that wealthy and educated Hong Kong immigrants paved the way for a larger group of Cantonese speaking immigrants such that the outcome is a Cantonese-speaking community that is socio-economically diverse. Furthermore, the wealth that middle and upper class immigrants injected into the Greater Toronto economy was very influential in changing the pre-1960s image of Chinese-Canadians as "inferior, second-class citizens" (Chan 2011:154). Wealth became even more influential in shaping the community in the 1980s. Two important events occurred during this decade. First was the ratification of the Sino-British Joint Declaration which set the details of Hong Kong's handover to the PRC government (see section 4.2.1). Second was the 1986 creation of a "business" class as a new class of immigrants under Canadian immigration law. This gave Hong Kongers from elite backgrounds, who had already benefitted from earlier changes, an even bigger advantage by having a separate class of immigrants for those with large monetary investments (Chan 2011:153). With a climate of fear about what would happen in Hong Kong after 1997, many wealthy business people from Hong Kong took advantage of this change in immigration policy. From 1986 to 1996, 47 percent of all business class immigrants to Canada came from Hong Kong (Chan 2011:154). The overall number of immigrants coming from Hong Kong also increased substantially during these years. In 1987, there were 16,170 immigrants from Hong Kong. This number almost tripled to 44,000 in 1994 (Chan 2011:154). Further contributing to the strong presence of Hong Kong Cantonese in Toronto during these years was the student population. From 1985 to 1995, the largest source of students entering Canada with visas was Hong Kong (Chan 2011:128).

Much of the growth in the Cantonese-speaking community in the 1980s and 1990s was in suburban areas of Toronto. The appeal of the suburbs was the wide availability of modern homes in new buildings built on large lots of land for relatively low prices. Such spacious homes would have otherwise been an unattainable luxury in Hong Kong (Chan 2011:155). Many wealthy Hong Kong investors also took advantage of the large tracts of open space to build shopping malls. Dragon Centre was the first of these major developments. Open in 1984, Dragon Centre set the model for other "Chinese malls" that would open up in several Toronto suburbs (Chan 2011:155–156). These Chinese malls became suburban "Chinatowns" and typically included office buildings, banks, banquet halls, restaurants, and other businesses all with Chinese signs. Cantonese became the primary language in these businesses. They were built with the automobile in mind and supported by the large population of Hong Kongers who purchased homes in the suburbs. The number of Chinese malls in the Greater Toronto Area (GTA) would grow to more than 40 by 2001 (Chan 2011:157). Some geographers have referred to these developments as "ethnoburbs" (Chan 2011:156). Ethnoburbs with the largest concentration of ethnic Chinese included Richmond Hill (34,615 out of 161,695 or 21.4%) and Markham (89,300 out of 261,573 or 34%). One of these malls, the Mississauga Chinese Centre, has even earned the distinction of being the only shopping center in all of Canada to be recognized as a tourist attraction because of its Chinese style architectural features (Chan 2011:158). Two of the Chinese malls in Markham are recognized as among the largest Asian themed malls in all of North America (Chan 2011:162).

The ethnoburbs are now home to the majority of ethnic Chinese in the GTA (Chan 2011:132). According to the 1996 census, only 9 percent of Chinese in the GTA live within the Municipality of Toronto. In contrast, Markham is home to 25 percent of the Chinese population in the GTA while Richmond Hill is home to 20 percent. Other suburbs with a large population of ethnic Chinese include Scarborough (home to 17 percent of the Chinese population of the GTA), East York (6% of the Chinese in the GTA), and Mississauga (5% of the Chinese in the GTA). A socio-geographical distinction has since emerged. While the Chinese population within the Municipality of Toronto is primarily working class, the Chinese population in the ethnoburbs is primarily middle-class and educated. In spite of these social class differences, these different Chinese communities within the GTA are not completely isolated from each other. For example, Chan mentions that many middle-class Chinese from the suburbs spend a lot of time in Toronto's Chinatown to shop for fresh produce at specialty stores and for weekend dim sum (2011:132).

The high overall socio-economic status of Hong Kong immigrants was crucial in facilitating maintenance of Cantonese. As Siemiatycki et al. have said, "Toronto's Chinese community is sufficiently large and affluent, in sections, to promote an impressive commercial, media, and marketing presence" (2003:408). Places of commerce, as has already been discussed, include multiple "Chinatowns" as well as suburban malls where Cantonese is widely used. Many of these businesses also advertise in local media. The importance of commerce and advertisement is in creating many opportunities outside of home where one can be exposed to Cantonese and where one can use Cantonese to communicate one's needs. Chinese mass media in Toronto created even more opportunities for use of Cantonese and exposure to Cantonese beyond one's household. Some of these newspapers such as *Sing Tao* and *Ming Pao* were North American subsidiaries of major Hong Kong newspapers (Chan 2011:189). In the 1990s, Toronto

Chinese community newspapers appeared including *New Star Weekly*, *Chinese Canadian Times*, and *Canadian Chinese News*. English-language magazines such as *Maclean's* even added Chinese editions to compete with the Chinese market. The *Markham Communicator*, named for the largely Chinese ethnoburb, became Canada's first Chinese-English bilingual newspaper. Perhaps even more important for Cantonese language maintenance are the presence of multiple TV and radio stations. This includes OMNI and Fairchild, a subsidiary of Hong Kong Television Broadcast Limited. Some local stations such as CFMT and CHIN are multilingual stations that include Cantonese programs. These stations include programs such as news, documentaries, and movies in Cantonese. Both Hong Kong born and Canadian born Cantonese speakers watch programs on these stations. According to Man (2006:221), 46.4% percent of young Cantonese-speaking Canadians (27% of whom are Canadian born) watch Cantonese TV very often; 26.8% watch often; 11.6% watch sometimes; and 10.7% watch occasionally. Only 4.5% never or very seldom watch Cantonese TV.

Heritage prestige is not unique to Cantonese in Toronto. As a whole, Toronto has been described as the "most multilingual city in the world". It is a city where 44% of the population speaks a mother tongue that is not an official language according to the 2016 Census[18]. City residents also view multilingualism and multiculturalism favorably. For example, some politicians and city planners have taken pride in the fact that Toronto has one of the highest percentages of foreign born and multilingual residents in North America (cf. Berridge 1995). In another example, a *Toronto Star* newspaper article described how different parts of the GTA are "a conurbation of neighbourhoods [sic], rather than ghettos" (Taylor 2007). Describing these

[18] https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=3520005&Geo2=PR&Code2=01&Data=Count&SearchText=toronto&SearchType=Begins&SearchPR=01&B1=Language&TABID=1

areas as not “ghettos” suggests the lack of a connection between ethnic minority identity and working-class status. The lack of negative social stigma may mean fewer reasons for second-generation Canadians of all ethnicities to want to completely assimilate to Anglophone Canadian culture.

Another way in which Cantonese language maintenance was supported was through Cantonese language instruction in schools. The loosening of immigration restrictions in the 1960s and 1970s also coincided with an era in which the Canadian government became more embracing of multiculturalism and the presence of diverse immigrant groups. In 1971, Canada became the first country in the world to adopt a multiculturalism policy. This was followed by the Multiculturalism Act of 1988 (Chan 2011:126). It was under such a climate that support for heritage language programs developed especially in the GTA. Ontario Regulation 154 requires heritage language instruction if “written requests are received from parents on behalf of 25 or more qualified persons of that school board or minority language section” (Man 2006:214). Heritage language classes are offered either after or on weekends. Cantonese language instruction has been one of the most successful heritage language programs. While some heritage language programs in Toronto have had declining enrollment, demand for Cantonese instruction has increased (Man 2006:227). Figures available for 1987 showed 3,625 students studying Cantonese in 40 schools and Mandarin in five as part of the program. These 3,625 students comprised about a third of all of the 10,000 total number of students enrolled in the Heritage Language Program. To accommodate increasing demand, many private language schools have also been established. In 1984, there were 10 such schools. The number increased by ten-fold to 100 by 1997. These years coincided with the years in which there was a huge growth of Hong Kong immigrants due to the ratification of Sino-British Joint Declaration. Once again, this highlights the influence of wealthy Hong Kong immigrants. Furthermore, the heritage prestige of Cantonese is enhanced by the desire to maintain Cantonese language skills because many children have learned to see future economic benefits in the ability to speak a variety of Chinese (Man 2006:223).

Overall, the combined influence of commerce, advertising, mass media, and heritage language instruction is a 70% of language retention rate among Canadian-born Cantonese heritage speakers. The most recent available census also shows 170,485 Cantonese speakers in the GTA. The only language with more speakers in the GTA is English. The second most widely reported variety of Chinese is Mandarin with 100,050 speakers. The census also reports 157,145 speakers of Chinese (not otherwise specified). This likely includes speakers of more than one variety of Chinese as well as many Cantonese speakers who chose the ambiguous term “Chinese” rather than “Cantonese”. Thus, the actual number of Cantonese speakers in the GTA is likely to be even higher than the 170,485 reported. From the 1960s through the 1990s, Hong Kong was the source of the largest number of ethnic Chinese immigrants. It was not until after the 1997 handover that the number of immigrants from Mainland China began to surpass the number of immigrants from Hong Kong (Chan 2011:168). The Mandarin speaking newcomers from the Mainland tended to be less economically successful than were the Hong Kong immigrants (Chan 2011:168). Thus, Cantonese has remained the dominant variety of Chinese in the GTA both in terms of numbers and in terms of the socio-economic status of its speakers.

Though not as numerous as the number of speakers of Cantonese, the relatively high number of speakers of Mandarin raises the question of whether or not there could be influence from Mandarin to Cantonese in Toronto. The hypothesis that there could be influence is based partially on the essentialist assumption that both groups are part of the same speech community because both groups share the same ethnic identity. Chinese ethnicity, however, poses problems to models of identity and sound change that assume isomorphism between identity and speech patterns. Cantonese and Mandarin are phonologically very different from each other and spoken by populations with distinct histories of migration to Canada. These groups also have different mass media viewing and listening preferences. For example, according to one survey conducted in 2008, the most popular radio station among Toronto Cantonese speakers is Fairchild (Chan 2011:189). For Mandarin speaking Torontonians, however, the most popular radio station is 680 News, which is actually an English station.

All of the GEN 1 immigrants that will be examined arrived in Canada before 1990. At this time period, barely half of Hong Kongers were able to speak and understand Mandarin. As in the case of Hong Kong, much of the transfer form Mandarin to Cantonese is likely to be transfer mediated by the Chinese writing system and hence loan words. If we expect Mandarin to have an influence on the direction of sound change in Toronto Cantonese, Mandarin would have to be spoken by a significant proportion of GEN 2 speakers. Evidence however shows that GEN 2 speakers are less likely to speak Mandarin than they are to speak English. GEN 2 speakers are also less literate in Standard Chinese than are GEN 1 speakers. These factors suggest few reasons for Mandarin to influence Cantonese except perhaps in loan words.

Instead, there is much stronger evidence for influence of English on Cantonese. This is because English is the language of instruction in the school system in Toronto. Thus, all children who are educated in the school system in Toronto learn English. The fact that knowledge of English is universal among GEN 2 Cantonese speakers, but Mandarin is not means that all of these speakers are capable of transferring features of English to Cantonese. This is already evident in the high use of code-switching with English in the recordings while examples of code-switching with Mandarin are extremely few. In a sense, this illustrates Bacon-Shone and Bolton's (1998) claim about the importance of English in the extended Hong Kong speech community. Just as code-switching with English is important in Hong Kong, it is also important in Toronto.

Previous research has shown a general lack of substrate influence in the English spoken by GEN 2 Cantonese-English bilinguals. For example, Hoffman and Walker (2010) show that GEN 2 ethnic Chinese, many of whom are also Cantonese speakers, show the same constraints on t-deletion and advanced variants of the Canadian Vowel Shift as ethnic Italians and those with British or Irish ancestry. The same constraint rankings, they argue, are evidence for a unified speech community. Although the actual rates of these variables do differ across ethnic groups, Hoffman and Walker (2010) interpret these different rates as evidence for individual speakers differing in how they express their ethnic identities.

The general lack of substrate influence on the English spoken by GEN 2 Cantonese-English bilinguals does not rule out the possibility that these speakers might transfer features of Toronto English to their Cantonese. With Toronto English pronunciation patterns part of their repertoire, it becomes possible for GEN 2 Cantonese-English bilinguals to transfer these patterns to the Cantonese they speak. Does this actually happen? This is the empirical question that will be addressed by the current project. Furthermore, if there is evidence of transfer from Toronto English, is the transfer via loan words or direct?

There is evidence that speakers within the community recognize different ways of speaking Cantonese. Nagy (2016), for example, cites an internet discussion board in which the poster says, "some of the accents are terrible, you can tell they're Canadian cantonese [sic] speakers (2016:21). Another poster from the same internet site says, "what bothers me, is that it's not authentic Cantonese, but *canadian cantonese* [sic]" (ibid.). Thus, in both of these posts, "Canadian Cantonese" is explicitly named and is evaluated in contrast to other forms of Cantonese. It is clear that some people notice something distinctive about the way Cantonese is spoken in Canada in contrast to the way it is spoken in Hong Kong. The linguistic features that they notice are uncertain. Differences in vowel production, the main topic of this dissertation, could be a possibility as has been shown to be distinguishing features in varieties of North American English (Labov, Ash, and Boberg 2006).

5.0 LINGUISTIC ISSUES, DATA, AND METHODS

In this chapter, I present the details of the current study. The specific research questions address inter-generational changes in the 11 surface monophthongs of Cantonese based on F1 and F2 measurements. These questions are spelled out as follows:

(Q1) Is there evidence for contact-induced inter-generational vowel shifting in native vocabulary?

(Q2) Is there evidence for contact-induced vowel mergers or vowel splits in native vocabulary? Four specific parts of the Cantonese vowel space are addressed:

(Q2a) Is there evidence for a merger between /y/ and /u/?

(Q2b) Is there evidence for an increasing acoustic split between /i/ and /ɪ/?

(Q2c) Is there evidence for an allophonic split in /ɛ/?

(Q2d) Is there evidence for an allophonic split in /ɔ/?

(Q3) To what extent can demographic (Sex, Age), ethnic orientation (overall EOQ score, individual EOQ responses, to be explained in Section 5.3.2), CAN % Score (see Section 5.3.3), CAN WC Score (see Section 5.3.3), and ENG WC score (see Section 5.3.3) account for the propagation of the specific shifts, mergers, and splits observed in the data?

The broader theoretical aim of this study is to address the following question:

(T1) What are the implications of the findings from this study for models of contact-induced sound change (cf. Labov 2007; Thomason and Kaufman 1988; van Coetsem 1988; 2000)?

The subsections that follow will discuss background information relevant to the current study. In 5.1, I discuss linguistic issues. This includes a discussion of the Cantonese vowel and tonal system. I present hypotheses about how contact induced change could lead to shifts, mergers, and splits in Section 5.2. In Section 5.3, I present details about the corpus and the data analyzed. This will be followed by a description of the methodology and data processing procedures in Section 5.4. I will conclude this chapter in Section 5.5 by discussing data analysis procedures.

5.1 CANTONESE VOWELS AND TONE

5.1.1 The Cantonese Vowel System

The Cantonese vowel inventory includes 11 surface monophthongs and 11 surface diphthongs. For comparison, the Toronto English inventory has 9 surface monophthongs (including schwa) and 7 surface diphthongs (with Canadian Raising variants included). Both languages, thus, have a typologically large number of acoustically distinct vowels[19]. Early bilingualism involving two genetically distinct languages (with one being non-Indo-European) that both have large vowel

---

[19] In the *World Atlas of Linguistic Structures*, Maddieson (2013) defines a large inventory as having 7 or more vowels.

systems is an important case study for addressing the extent to which models of sound change and contact induced change hold across different types of contact settings and across different languages.

To limit the scope of the current study, I focus only on monophthongs and only on vowel quality as measured in terms of midpoint F1 and F2. Future research could also consider other measurements such as vowel duration, changes in F1 and F2 over time, and Euclidean Distance. Such measurements would address other ways in which these vowels could be changing. For example, could some monophthongs be becoming diphthongal or could there be changes in vowel duration (and by extension vowel length as a phonological feature)? These are all worthwhile questions to pursue, but for the purpose of the current study, I consider only three types of change that can be measured in terms of midpoint F1 and F2: shifts (Q1), mergers (Q2), and splits (Q2).

To define the vowel categories analyzed, I follow the description in Zee (1999), which recognizes 11 surface monophthongs. Seven of these are described in the Cantonese literature as tense or long (Table 11), while the other four are described as lax or short (Table 12). In the second column in both Table 11 and Table 12 is the Jyutping Romanization of each vowel. In the third column in both of these tables is a list of environments in which each of these 11 vowels can occur. These environments are described in terms of coda consonants that can follow each vowel. A "#" indicates open syllable context. All the sample words shown are pronounced with a high-level tone (˥ in IPA). A visual representation of these vowels is shown in Figure 2.

Table 11. Cantonese tense (or long) monophthongs (following Zee 1999)

Vowel in IPA	Jyutping Transcription	Environments	Example	Gloss
i	<i>	#, _p, _t, _m, _n	si˥	‘silk’
y	<yu>	#, _t, _n	sy˥	‘book’
ɛ	<e>	#, _p, _t, _k, _m, _n, _ŋ	sɛ˥	‘to lend’
œ	<oe>	#, _k, _ŋ	hœ˥	‘boot’
a	<aa>	#, _p, _t, _k, _m, _n, _ŋ	sa˥	‘sand’
ɔ	<o>	#, _t, _k, _n, _ŋ	sɔ˥	‘comb’
u	<u>	#, _t, _n	fu˥	‘husband’

Table 12. Cantonese lax (or short) monophthongs (following Zee 1999)

Vowel in IPA	Jyutping Transcription	Environments	Example	Gloss
ɪ	<ik>/<ing>	_k, _ŋ	sɪk˥	‘color’
ɵ	<eo>	-t, _n	sɵt˥	‘shirt’
ɐ	<a>	_p, _t, _k, _m, _n, _ŋ	sɐp˥	‘wet’
ʊ	<uk>/<ung>	_k, _ŋ	sʊk˥	‘uncle’

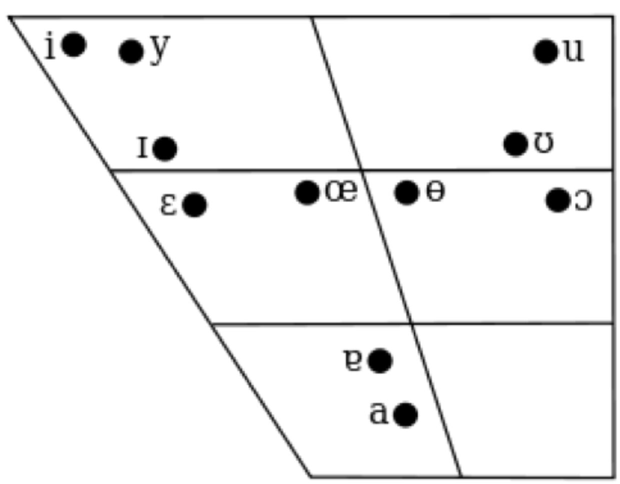

Figure 2. Cantonese vowel space with 11 monophthongs

Also important to understanding the Cantonese vowel system is how each vowel fits in to the syllable template. Bauer and Benedict (1997:315–316) formulate the template as follows:

$$\overset{T}{(C_i)\ N\ (C_f)}, \ N = S \text{ or } V$$

At the minimum, all Cantonese syllables must have a tone (T) and a nucleus (N). The N can be either a syllabic nasal (S) or a vowel nucleus (V), but not both. Syllables that contain a V can optionally have an onset consonant (C_i), coda consonant (C_f), or both an onset and a coda. Only one segment can occupy either the C_i or the C_f slot. Thus, neither onset clusters nor coda clusters are possible in Cantonese[20]. One stipulation to this description is that the [kʷ] in words like [kʷa55][21] ('melon') or the [kʷʰ] in words like [kʷʰɐn21] ('skirt') must be treated as single segments rather than sequences of a stop followed by a glide. Any consonant in the Cantonese inventory can occur as a C_i but for C_f, the possible consonants are limited to nasals (/m/, /n/, /ŋ/), stops (/p/, /t/, /k/), and glides (/j/, /w/, /y/). If a glide occurs as a C_f, the syllable contains a diphthong. This accounts for the fact that Cantonese diphthongs occur only in open syllable context while syllables closed by a coda consonant can only have monophthong nuclei.

The description of the Cantonese vowel system that I adopt in the current study is a surface-based description. The alternative would be a description based on abstract groupings of surface vowels into phoneme categories. Unlike an abstractionist description, the advantage of a surface-based description is that it does not require making a priori assumptions about which allophones belong to which phoneme categories. This, in fact, has been a very controversial topic

[20] Some speakers do have clusters in their speech, but for those that do clusters are limited to onomatopoeic expressions, loan words, and in contracted forms of multi-syllabic words in rapid speech (Bauer and Benedict 1997:319).

[21] See Table 14 in Section 5.1.2 for a list of Tone Numbers.

in Cantonese phonology. In contrast, there has been relatively little dispute about how many surface vowels there are and what distributional patterns they have.

Distributional patterns related to the environmental contexts in which each vowel occurs are a central part of Cantonese vowel phonology. For example, one pattern is based on open vs. closed syllable environment. The tense vowels shown in Table 11 are the only monophthongs that can occur in open syllable context. The lax vowels ([ɪ], [ɵ], [ɐ], [ʊ]) as shown in Table 12, on the other hand, only occur in closed syllable environment. Among the lax vowels, [ɐ] can occur in the most environments. [ɪ], [ɵ], and [ʊ] occur in restricted environments based on the place of articulation of the following segment. The lax vowels are also the vowels that have most often been transcribed using different IPA symbols. For example, alternative transcriptions for [ɪ], [ʊ], [ɐ], and [ɵ] include [e], [o], [ʌ], and [ø] respectively.

Other distributional patterns relate to complementary distribution relationships. In abstractionist approaches, this has been the source of many of the debates about the Cantonese vowel system. This is because of multiple complementary distribution relationships that make multiple possible groupings of phones into phoneme categories. To illustrate, Table 13 shows possible rimes from three pairs of vowels. Each column includes a possible coda while each row includes a vowel. As shown, [ɵ] and [œ] occur in complementary distribution with [ɵ] occurring only before alveolar segments (ex: [_ɵn] and [_ɵt]) and as part of the diphthong [ɵy] while [œ] occurs only preceding a velar consonant (ex: [_œŋ] and [_œk]) or in open syllable context. Since both [ɵ] and [œ] are mid round vowels, we can conclude that these sounds are allophones of the same phoneme. Yet, there is another possible analysis. [ɵ] also occurs in complementary distribution with [ʊ]. These two vowels are also round and relatively close to each other in the vowel space. Hence, they share some phonetic similarity and can also be grouped together as allophones of the same phoneme.

[ʊ], however, can also be grouped together with [u] in a complementary distribution relationship since the former occurs only before velar consonants and the latter occurs elsewhere. Once again, there is phonetic similarity. In this case, both are high, back, and round. Thus, [ʊ] and [u] can also be grouped together as allophones of the same phoneme based on these two criteria (complementary distribution and phonetic similarity). With [ɪ] and [i], there is a parallel pattern with the former occurring only before velar consonants and the latter occurring elsewhere. These two vowels can be seen as the front, unrounded counterparts of [ʊ] and [u] and can also be grouped together as allophones of the same phoneme based on the same criteria.

Table 13. Rime group table showing complementary distribution of three pairs of vowels

	#	y	m	n	ŋ	p	t	k
ɵ		_ɵy		_ɵn			_ɵt	
œ	_œ				_œŋ			_œk
ʊ					_ʊŋ			_ʊk
u				_un			-ut	
ɪ					_ɪŋ			_ɪk
i			_im	_in		_ip	_it	

Even more analyses are possible if one considers vowel length and the possibility of grouping monophthong and diphthong phones together. Bauer and Benedict (1997), for example, have challenged analyses that group [i] and [ɪ] together and [u] and [ʊ] together. They consider acoustic studies showing that the distinction between [ɐ] and [a] is based primarily on length rather than on vowel quality. Similarly, they also discuss studies of both [ɪ] and [ʊ] that show that these two vowels are acoustically "short diphthongs". Since duration appears as a contrastive feature for multiple pairs of vowels, they propose an analysis that groups all Cantonese vowels based on length as the primary feature. This leads to the conclusion that [ɪ] and [ʊ], which they propose transcribing as [eʲ] and [oʷ], should be grouped together with the diphthongs [ej] and [ow] rather than with [i] and [u] as in many other proposals.

As illustrated by this discussion, the problem with adopting an abstractionist description of the vowel system is that doing so would require choosing one analysis out of many possible analyses including many that have not been discussed above due to space considerations. In fact, at least 21 different analyses[22] have been proposed in the literature. The number of vowel phonemes proposed has ranged from five to 11 (Yip 1996)[23]. Another problem with many abstractionist descriptions is that they assume that length is a distinctive feature. Obtaining vowel duration measurements to support this assumption, however, is not part of the current study. With a surface-based description, however, there is no need to make a priori assumptions about which vowel phones are allophones of which vowel phonemes.

There is also evidence that surface-based descriptions of Cantonese do not deviate substantially from underlying representations of Cantonese phonology. Yip notes that "the development of phonological theory has been largely driven by languages with alternations,

[22] I counted this number based on several sources. Yue-Hashimoto (1972) presents a critical review of nine different proposals and concludes by presenting a tenth proposal. Bauer and Benedict (1997) also present a critical review of various proposals including some also discussed by Yue-Hashimoto (1972). The total in these two sources combined is 17. Barrie (2003) is a more recent analysis not included in these two sources. Several Romanization systems also implicitly adopt different analyses not represented in these other proposals including the Sydney Lau System, the Yale System, and the Jyutping System. Thus, 17 plus these four make 21 different analyses.

[23] If we include surface-based analyses that treat diphthongs as derived from monophthongs, the total goes up to 13. The two extra vowels would be vowels that occur only as part of diphthongs while the other diphthongs are derived from all of the other 11 vowels. If we treat diphthongs as not derived from monophthongs, however, the total goes up to 22.

where considerations of lexical economy make the postulation of abstract underlying forms and productive rules which transform these into surface forms very natural (Yip 1996:1)." Languages like Cantonese, Mandarin, and other varieties of Sinitic, however, are languages that lack alternations. For this reason, Yip says that such languages "have never fitted smoothly into such theories (ibid.)" She argues as a universal principle that learners of such languages "will naturally internalize the forms closest to the surface, absent paradigm pressure to do otherwise" (ibid.). Thus, since Cantonese lacks morpho-phonological alternations, there is little motivation for Cantonese learners to posit underlying forms that are more abstract than the surface forms that they hear. According to this view, the 11 surface monophthongs present in Cantonese transparently map on to 11 vowel phonemes. Yu (2000) supports this idea through a perceptual study showing that Cantonese speakers recognize more distinctions than what would otherwise be expected if phoneme categories were based simply on the most parsimonious accounts of complementary distribution patterns. Parsimony in the number of phonemes, however, does not necessarily reflect the number of distinct phonological units that speakers recognize.

Thus, for the purpose of the current dissertation, I adopt a surface-based description of the Cantonese vowel system. This follows a relatively uncontroversial issue in Cantonese linguistics. Descriptions of Cantonese vowel phonetics and phonology are in unanimous agreement about the lack of finer level quality distinctions beyond the 11 surface categories[24]. Thus, it is not expected that there would be any conditioning effects on these 11 vowels in terms of vowel quality[25] with the exception of two allophonic splits, which I discuss in Section 5.2.

[24] Zee (2003) has recognized finer distinctions based on duration. The phonological significance of these distinctions, however, is unclear.

[25] One possible exception is with /ɐ/. Bauer and Benedict (1997), however, present conflicting details about this vowel. On one page, they note that /ɐ/ is retracted when preceding labial consonants (Bauer and Benedict

Furthermore, the surface-based approach is similar to the use of lexical classes in variationist studies of English. For instance, KIT and PIN are sometimes treated as different lexical classes in studies of English variation. Even though these two words are pronounced with the same vowel in many dialects, treating them as distinct word classes leaves open the possibility that there may be low-level phonetic differences conditioned by the following consonant in some dialects. Thus, contra (Tse 2016a; 2016b), the vowels [ɪ] and [ʊ] will be treated as separate vowel categories rather than as allophones of /i/ and /u/ respectively.

5.1.2 Tone

Since Cantonese is a tonal language, tone could be a phonological factor conditioning vowel variation and change. Thus, tone is related to Q2 and to Q3. Unlike some varieties of Sinitic such as Mandarin, Cantonese does not have any toneless syllables. Each Cantonese syllable must have a tone category. Table 14 includes a list of all 6 contrastive tones along with a verbal description of the associated contour pattern and the Chao Tone Number equivalent. While the Tone number in the first column is an arbitrary number, the Chao Tone Number is a perceptual scale from 1 (lowest) to 5 (highest) that shows the relative height contour. For example, “35” indicates mid rising while “21” indicates low falling and “55” indicates high level. From this point forward, I will use the Chao Tone Number in IPA transcriptions. In Jyutping transcriptions, I will use the Jyutping tone category. For example, the word for poetry in IPA would be [si55] while in Jyutiping it would be <si1>.

1997:71). Later on, however, they discuss a set of words that may have been formerly pronounced as [_ɔp]/[_ɔm] that are now merged with words pronounced as [_ɐp]/ [_ɐm] (Bauer and Benedict 1997:419–420). This suggests a phonological contrast rather than phonological conditioning.

There has not been any research (at least not known to the author) showing tone categories conditioning vowel variation in Cantonese. Distributional patterns that involve specific tone categories occurring with specific syllable types and with specific vowels, however, are well documented. Some of these co-occurrence restrictions are related to the diachronic development of tonal distinctions in Cantonese while others are accidental. One pattern involves a distinction between checked and non-checked syllables. Checked syllables include syllables that end with a coda stop (/p/, /t/, and /k/) while all other syllable types including open syllables and syllables closed with a nasal consonant are non-checked. It is only in non-checked syllables that contrasts across all six tone categories are possible. In checked syllables, however, only four of these tones are possible. Three of these tones are level tones. In older analyses of Cantonese, the level tones occurring in checked syllables were analyzed as distinct tone categories. This resulted in a total of nine different tone categories. In more recent analyses, however, these level tones in checked syllables are treated as the same level tones that occur in non-checked syllables. These tones in checked syllables, however, are phonetically shorter in duration. To indicate the relative shortness of these tones, some scholars have indicated these tones using only a single Chao Tone number.

Checked syllables also have further co-occurrence restrictions based on whether the vowel is tense or lax. Tone 1 can occur only with lax vowels while Tone 3 occurs only with tense vowels. Tone 6 can have either lax or tense vowels. There are only a handful of exceptions to these co-occurrence restrictions. Most of these exceptions involve words that are either not inherited from Middle Chinese or words that have alternate pronunciations in the literary register of Cantonese (Yue-Hashimoto 1972:177). Thus, the pattern of Tone 1 vs. Tone 3 mapping onto a lax vs. tense distinction appears to have been an outcome of the diachronic development of Cantonese phonology.

Table 14. Cantonese tones

Jyutping Tone Category	Description (With Chao Tone Numbers)	Non-Checked Syllable Examples	Checked Syllable Examples	
			Lax Vowels	Tense Vowels
1	High level (55)	/si55/, ‘poetry’	/sɪk55/ → [sɪk5], ‘to know’	--
2	Mid rising (35)	/si35/, ‘history’	/jɐn21 mat22/ → [jɐn21 mɐt35], ‘character’ (‘person + thing’)	/ŋa21 t͡sʰat33/ → [ŋa21 t͡sʰat35], ‘toothbrush’ (‘tooth + brush’)
3	Mid level (33)	/si33/, ‘to try’	--	/sɛk33/ → [sɛk3], ‘to kiss’
4	Low falling (21)	/si21/, ‘time’	--	--
5	Low rising (23)	/si23/, ‘city’	--	--
6	Low level (22)	/si22/, ‘matter’	/sɪk22/ → [sɪk2], ‘to eat’	/sɛk22/ → [sɛk2], ‘stone’

Tone 2 is also possible with both lax and tense vowels in checked syllables but occurs only in morphologically derived forms that surface through reduplication or through a morphological process known in Cantonese linguistics as <bin3 jam1> (“change tone”). The bin3 jam1 (also spelled “pinjam” as in Yu 2007) process involves changing the underlying tone of the second syllable to a mid-rising tone in the formation of certain compound words. The two examples of Tone 2 (or “35” in the Chao System) in Table 14 are both derived from bin3 jam1[26]. To illustrate, the word [mɐt22] (<mat6>) on its own means ‘thing’. When this word is combined with the word for ‘person’ to form a compound word, the 22 tone in /mɐt22/ changes to a 35

[26] Lee (2014) has analyzed the bin3 jam1 tone as a floating tone similar to a tonal morpheme found in some African languages.

tone. Similarly, the word /t͡sʰat33/ (<caat3>) has a 33 tone on its own and means ‘brush’. In the compound word for ‘toothbrush’, [ŋa21 t͡sʰat35] (<ngaa4 caat2>), the tone changes to a mid-rising tone.

With all of these co-occurrence restrictions, it may be difficult to completely separate tonal conditioning effects from the conditioning effects of adjacent segments. Ultimately, what matters for the current study is whether or not there are inter-generational differences in phonological conditioning rather than what the exact phonological factors are.

5.2 HYPOTHESIZED VOWEL SHIFTS, MERGERS, AND SPLITS

In this section, I discuss a set of hypothesized vowel shifts, mergers, and splits that may be found in the data. Since this dissertation is only one study, the hypothesized changes addressed will be limited to eleven possible vowel shifts, one possible merger, and two possible splits. The specific hypotheses are presented in each of the following sub-sections.

5.2.1 Hypothesized Vowel Shifts (Q1)

To address the first research question about whether or not there is evidence for contact-induced vowel shifts, we would first need to establish whether or not there is inter-generational change. Lack of change would simply be evidence for inter-generational phonological maintenance and transmission. If there is change, however, the interpretation of what the change means would depend on the direction of change. In some cases, both internal motivation and contact-induced change would predict the same direction of change. For such cases, Homeland data would be considered to determine whether or not the same change is taking place in the Homeland. The lack of the same change in the Homeland variety would strengthen an explanation based on contact-induced change.

In Table 15 and Table 16 below, I present two sets of hypotheses about how each vowel category could change due to either internal motivation or due to assimilation with the most phonetically similar Toronto English vowels. A third possibility, following Flege (1995) and discussed in Chapter 3, is dissimilation (or deflection) from the most phonetically similar Toronto English vowel. This possibility will not be considered because it is difficult to evaluate without normalized acoustic data from the English spoken by the same population of speakers. Further complicating matters is that dissimilation could go in multiple directions.

In Table 15 are the tense vowels while in Table 16 are the lax vowels. In both of these tables, the column labeled "internal" shows the expected direction of change following the Principles of Chain Shifting while the column labeled "Toronto English Assimilation" shows the direction of change expected if Cantonese vowels shifted to the most phonetically similar vowel in Toronto English.

Labov's (1994) Principles of Chain Shift are as follows: (i) long vowels rise, (ii) short vowels fall, and (iii) back vowels move to the front. Cantonese has seven tense (or long[27] vowels). According to Principle I, all of the long vowels in Cantonese would rise. Some of these long vowels, however, are also back vowels. According to the third principle, these back vowels, which include /ɔ/ and /u/, would front. Both fronting and raising are indicated in the internal

[27] For the purpose of the current study "tense" and "long" are treated as synonymous. "Lax" and "short" are also treated as synonymous.

motivation column for these two vowels. Cantonese also has four lax (or short) vowels. According to Principle II, these vowels will lower. This is indicated in Table 16.

Table 15. Possible vowel changes (tense vowels)

Vowel	Internally Motivated Change	Toronto English Assimilation
I	raising[28]	lowering
Y	fronting/raising	retraction
U	fronting/raising	fronting
Œ	raising	F3 changes
ɛ	raising[29]	lowering
ɔ	fronting/raising	lowering
A	raising	fronting or retraction

Table 16. Possible vowel changes (lax vowels)

Vowel	Internally Motivated Change	Toronto English Assimilation
ɪ	lowering	lowering/retraction
ʊ	lowering	lowering
ɵ	lowering	Raising
ɐ	lowering	--

[28] Diphthongization of /i/ to /ej/ is another possibility. Bauer and Benedict (1997:334–335) note that some Cantonese words now pronounced with the diphthong [ej], were pronounced instead with the monophthong /i/ at the turn of the 20th century. This is still the case in some Yue dialects. In modern Hong Kong Cantonese, the demonstrative pronoun is variably pronounced as [ni]/[nej]/[li]/[lej]. In this case, the variation appears to involve lexical diffusion. Since this involves diphthongization, I do not consider such variants in the current study.

[29] Readers familiar with English phonology may recognize /ɛ/ as a lax vowel. In Cantonese, however, it is a tense vowel.

Table 17. Possible vowel changes (tense vowels)

Cantonese Vowel	Loan word Examples with English source word in parenthesis from Bauer and Benedict (1997:383–394)	Toronto English Correspondences
i	[t͡ʃ**i**55.si35] (‘cheese’); [tʰ**i**p55.si35] (‘tips’); [kit33.tʰa55] (‘guitar’);	/i/, /ɪ/
y	--[30]	--
u	[m**u**55.fi35] ‘movie’; [gʊk55.k**u**35] (‘cocoa’); [kʰ**u**55.sɐn35] (‘cushion’)	/u/, /ow/
œ	[pʰ**œ**55.sɛn55] (‘percent’), [s**œ**21] (‘sir’)	/ɚ/
ɛ	[t**ɛ**55.ti21]‘daddy’; [t͡sʰ**ɛ**55.low35] (‘cello’), [t͡s**ɛ**55.si22] (‘jersey’); [mɛ55.t͡sa35] (‘major’)	/ɛ/, /æ/
ɔ	[p**ɔ**55] (‘ball’); [**ɔ**55.ta35] (‘order’), ‘toast’ [d**ɔ**55.si22]	/ɑ/, /ow/
a	[pʰ**a**55.si35] (‘pass’), [p**a**55.si35] (‘bus’), [t**a**55.l**a**35] (‘dollar’);	/æ/, /ɑ/

Table 18. Possible vowel changes (lax vowels)

Vowel	Examples	Toronto English Correspondences
ɪ	[tʰ**ɪ**k55] (‘tick’); [kʰ**ɪ**ŋ55] (‘king’); [kʰ**ɪ**k55] (‘cake’); [t͡s**ɪ**k] (‘jack’); [t**ɪ**k55.si21.kow55] (‘disco’)	/ɪ/, /ej/, /æ/
ʊ	[g**ʊ**k55.ku35] (‘cocoa’); [p**ʊ**k55.kʰa35] (‘broker’), [kʰ**ʊ**k55.ki21] (‘cookie’)	/ʊ/, /ow/
ɵ	[ka55.l**ɵ**n21] (‘gallon’); [sa55.l**ɵ**t22] (‘salad’); [s**ɵ**t55] (‘shirt’)	/ɪ/, /əɹ/
ɐ	[f**ɐ**n55] (‘fun’), [n**ɐ**m55.pa22.w**ɐ**n55] (‘number one’)	/ʌ/

[30] The only Cantonese loan word with a /y/ identified by Bauer & Benedict (1997) is [t͡ʃy.ku.lɪk] (‘chocolate’). I have excluded this word in the table because of uncertainty about whether or not it really is a loan borrowed directly from English. I suspect Portuguese is a possible source. The European Portuguese form is [ʃukuˈlat], which would better explain why high round vowels occur in the Cantonese form than an English source. Other dialects of Chinese (which may have borrowed the word from English, Portuguese, or another language) may be other possible sources (Patrick Chew, p.c.).

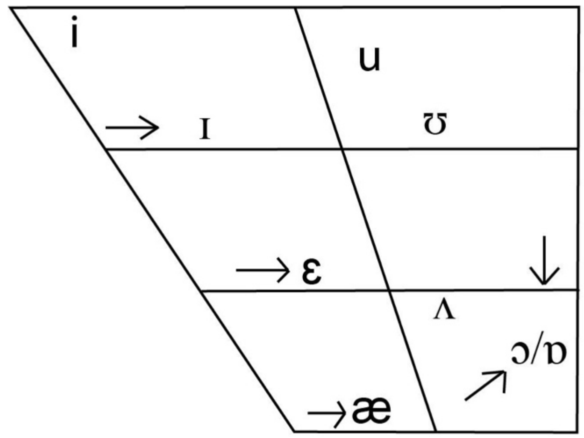

Figure 3. Toronto English monophthongs with arrows indicating the Canadian Vowel Shift

The third step in determining possible directions of change based on cross-linguistic assimilation was to compare the Cantonese vowel with the corresponding vowel in Toronto English. In Figure 3 is a vowel chart showing the eight monophthongs of Toronto English. Some of these vowels are involved with the Canadian Vowel Shift. The direction of movement of these vowels are indicated with arrows. Clarke, Elms, and Youssef (1995) describe the pivot of these changes as the merger between the low back vowels /ɔ/ and /ɒ/. According to their argument, this merger creates room for the lowering and/or retraction of three front lax vowels: /ɪ/, /ɛ/, and /æ/. The lowering and/or retraction of these vowels also follows Labov's (1994) Principle II. They result in a set of vowels that are lower and more retracted than in many other varieties of English.

For Cantonese /ɐ/, there is no a priori reason to suspect that the corresponding Toronto English vowel, /ʌ/ is significantly different along the F1/F2 space. Although the transcriptions are different, Cantonese /ɐ/ corresponds very consistently with Canadian English /ʌ/ based on correspondences[31] identified in loan words. For both of the high round vowels /y/ and /u/, the closest Toronto English counterpart is /u/. Like many dialects of English, Toronto English /u/ is fronted. In fact, its position in the F1/F2 space may actually be midway between Cantonese /u/ and /y/. Thus, one possibility is that Cantonese /u/ could be fronted under influence from Toronto English while the other possibility could be that /y/ is retracted.

For other cases, there are multiple correspondences. Often these depend on adjacent segments in English as well as Cantonese phonotactics. For example, Cantonese /u/ corresponds with Toronto English /u/, /ʊ/, and /ow/. Cantonese /ɪ/ also has three corresponding vowels in Toronto English: /ɪ/, /æ/, and /ej/. The /ej/ correspondence is an example of influence from Cantonese phonotactics. This sound correspondence occurs in the word for 'cake'. Since diphthongs occur only in open syllables in Cantonese, a monophthong occurs instead in the Cantonese loan. In all these corresponding English forms, however, the corresponding vowel is either identical (at least in transcription terms) or articulatorily lower. The same is the case for Cantonese /i/, which corresponds with Toronto English /i/ and /ɪ/, and for Cantonese /ɛ/, which corresponds to either Toronto English /ɛ/ or /æ/. Thus, the hypothesized change for /i/, /ɪ/, /ʊ/, and /ɛ/ is lowering since all four of these vowels have correspondences with lower vowels in Toronto English.

For vowels with corresponding Toronto English forms that correspond with forms involved in the Canadian Vowel Shift, we can predict that assimilatory influence would lead these vowels to follow the same direction of movement as in the Canadian Vowel Shift. For example, /ɛ/ is retracted according to the Canadian Vowel Shift. Thus, we might expect

[31] A few loan words identified by Bauer and Benedict (1997) have this vowel corresponding with /a/.

Cantonese /ɛ/ to also be retracted. This is the same direction of movement as expected based on correspondences with English /æ/. The low back merger also means that the closest counterpart to Cantonese /ɔ/ would be articulatorily lower in the vowel space. Cantonese /a/ is the vowel that is the most difficult to predict in terms of how it could assimilate to the most phonetically similar Toronto English counterpart. It is a low central vowel and it is likely intermediate between the retracted /æ/ in Canadian English and the /ɑ/ (retracted as part of the low back merger). Cantonese could either front towards /æ/ or retract towards /ɑ/.

In some cases, assimilatory influence would lead to the same expected outcome as internal motivation. For example, Toronto English /u/ is fronted. Assimilatory influence on Cantonese /u/ would also lead to the fronting of /u/ in Cantonese. This is also a shift that follows Principle III of the Principles of Chain Shifts. Some other examples include the lowering of /ɪ/. The lowering of /ɪ/ would follow Principle II of Chain Shifting, but it could also be due to assimilatory influence with Toronto English /ɪ/, which is also lowered due to the same principle.

The most problematic and least ambiguous cases for the TD Model would be cases involving a vowel shift influenced by Toronto English that is also opposite of the direction expected from chain shifting principles. An example of this would be the retraction of /y/. None of the principles of chain shifting would predict the retraction of this vowel. Given the relatively fronted position of Toronto English /u/, however, retraction of Cantonese /y/ could be due to assimilatory influence with the fronted Toronto English /u/. For the case of Cantonese /ɛ/, the phonological status of this vowel may play a crucial role. While Cantonese /ɛ/ is a tense vowel, Toronto English /ɛ/ is lax and is also the most phonetically similar vowel counterpart. Principle I could apply because Cantonese /ɛ/ is tense and would result in the raising of this vowel. Yet, assimilatory influence could also apply and result in a shift going in the opposite direction.

5.2.2 Evidence for Vowel Mergers (Q2a)

The second research question is about evidence for contact-induced phonological mergers or splits. Four specific hypotheses are considered. The first is a merger between /y/ and /u/. This part of the vowel system seems susceptible to influence from Toronto English because Toronto English, like most dialects of English, lacks a phonemic contrast between two high round tense vowels. Furthermore, the phonetic status of the high round vowel, /u/, in Toronto English is fronted (as in many other English dialects). The F2 of fronted /u/ could potentially overlap with the F2 range of /y/ in Cantonese and in many only languages. Could influence from Toronto English lead to a loss of distinction between /y/ and /u/ in Cantonese?

Chang et al.'s (2011) study, discussed in Chapter 3, addressed a similar contrast among heritage speakers of Mandarin. This study found no loss in phonological contrast. Instead, the results from this study showed that some heritage speakers may even retract /u/ as a way of maximizing cross-linguistic contrast. This study, however, was based on experimental data. Can the same results be replicated for a study of heritage Cantonese speakers using spontaneous speech data?

One important consideration in the sociophonetic study of vowel mergers is that vowel mergers in progress are often partial. There are several ways of addressing merger between /y/ and /u/. One way would be to determine if there is both inter-generational retraction of /y/ and inter-generational fronting of /u/. The simultaneous presence of both types of changes would support a change towards merger.

Another method that has become increasingly common in sociolinguistic studies of vowel merger is the "Pillai Score", a name given by Hay, Warren, and Drager (2006) to refer to the "Pillai-Bartlett statistic", which is one of the four common MANOVA (Multivariate Analysis of Variation) tests. Hay et al. describe the Pillai Score as a "summary of the degree to which two distributions are kept distinct" (2006:467). A Pillai score ranges from 0 to 1, with a higher score indicating greater F1 and/or F2 distance and a lower score indicating a more of a merger. Pillai scores can be calculated for individual speakers by using statistical software (R, SPSS, etc) to run a MANOVA model with F1 and F2 as the two dependent variables and vowel class as the independent factor for all tokens of the two vowel classes in question.

For the current study, the Pillai score seems better suited than other measures of vowel distinction and overlap such as Euclidean Distance, Mixed Effects Regression with Adjusted Euclidean Distance, and Spectral Overlap (see Nycz and Hall-Lew 2015 for a review) because Pillai scores model both F1 and F2 variation simultaneously unlike Euclidean Distance and Adjusted Euclidean Distance. Furthermore, in calculating a score for each individual speaker that measures overall difference, Pillai Scores can be more easily used to show how different speakers compare to each other in continuous terms. For instance, Speaker A with a Pillai score of 0.300 can be described as more merged for these two vowels than Speaker B who has a Pillai score of 0.500 while Speaker C with a Pillai score of 0.900 can be described as having two very distinct vowels. Pillai scores can also subsequently be used as input for statistical tests that include social variables. For example, Hall-Lew (2009) correlated speaker Pillai scores with age to show that younger speakers had lower Pillai scores for the vowels in LOT vs. THOUGHT in San Francisco English. The statistical significance of this correlation, thus, indicated an apparent time change towards merger of these two vowels. It is important to note that the Pillai score does "not provide statistical discrimination between those with near-merger and those with complete merger" (Hall-Lew 2010:3). As a point of reference, however, Hall-Lew (2009:150) arbitrarily chose a Pillai score of 0.200 and below as definitely merged (and 0.300 and below as a less conservative cut-off point for merger) in her study of the LOT-THOUGHT merger in San Francisco English.

5.2.3 Evidence for Increasing Split Between [i] and [ɪ] (Q2b)

Previous work showed that GEN 2 Toronto speakers have increased the acoustic distance between /i/ and /ɪ/ (Tse 2016b). This study was inconclusive about whether this changed is internally motivated or due to contact with Toronto English. The retraction or lowering of /ɪ/, for example would follow Labov's Principle II of Vowel Chain Shifting (short vowels fall). Toronto English also has a retracted /ɪ/ compared to many other dialects of English. Since both internal motivation and external motivation would predict the same outcome, it is not certain whether or not this change is a contact-induced change.

In this dissertation study, I include a much larger group of speakers and a larger number of tokens from each speaker to determine whether or not the same pattern holds. I will also use Pillai Scores to determine which specific speakers lead in having the greatest acoustic difference between these two vowels. To provide further evidence to support or refute a contact-induced change explanation, apparent time data from Hong Kong will be included in the analysis. The presence of the same change in Hong Kong would provide further support for internal motivation while its absence would support a contact-induced change account.

5.2.4 Evidence for an Allophonic Split in /ɛ/ (Q2c)

A previous study showed that GEN 2 speakers have developed a fronted allophone of /ɛ/ in pre-velar context (Tse 2016a). Although some dialects of Canadian English have a raised allophone of /æ/ in pre-velar context, this has been described as primarily a Western Canadian feature rather than a Toronto feature (Boberg 2008). One possible explanation I discussed is that this split may have developed to enhance the contrast between /i/ in pre-velar environment and /ɛ/ in pre-velar environment.

The problem with this previous study is that it was based on only 5 tokens of /ɛ/ in pre-velar context and 10 tokens of /ɛ/ in open syllable context for each speaker. Pre-velar context included both /ŋ/ and /k/. With a larger set of data for this dissertation study from a larger group of speakers, I will address whether or not this pattern holds. Pillai Scores will also be used to determine which specific speakers may be leading in this change. I also present a new hypothesis about how /ɛ/ could split under contact with Toronto English.

Under my new hypothesis, the tense/lax status of /ɛ/ is important. In Cantonese, /ɛ/ is categorized as a tense vowel (see discussion in Section 5.1.1), but in Toronto English it is lax. Cantonese /ɛ/ can also occur in both open and closed syllable contexts while Toronto English /ɛ/ occurs only in closed syllables. Toronto English also has a raised tense variant of the vowel /æ/ that occurs in pre-nasal context (Boberg 2008). I predict that /ɛ/ in Cantonese may be fronted in pre-nasal context due to phonetic similarity with the raised tense /æ/ in Toronto English. Since /ŋ/ is nasal, this would be consistent with my earlier study (Tse 2016a). /k/, however, is not nasal. In pre-/k/ environment, I predict retraction because /ɛ/ before coda stops may be more phonetically similar to the retracted /ɛ/ in Toronto English, since it occurs only in closed syllables. Thus, instead of the hypothesis presented in my earlier study (Tse 2016a), it is pre-nasal and pre-coda stop environment rather than pre-velar context that condition the splitting of Cantonese /ɛ/.

5.2.5 Evidence for an Allophonic Split in /ɔ/ (Q2d)

My earlier study also showed evidence of an allophonic split in /ɔ/ which involved articulatorily lowered variants in pre-velar context (Tse 2016a). The explanation I offered was that it is a change that enhances the contrast between /ʊ/ and /ɔ/ in pre-velar environment. Again, this was based on a small sample of 10 tokens in open syllable environment and only 5 in pre-velar (including both /k/ and /ŋ/). For this dissertation study, I will analyze a larger set of data including more phonetic environments to determine if this pattern holds. I will also use Pillai Scores to determine which specific speakers may be leading in this change if it is also found across a larger set of data.

5.3 THE DATA

5.3.1 The Corpus

The data for the current study comes from the HerLD corpus, a product of the Heritage Language Variation and Change (HLVC) in Toronto Project. This corpus includes hour-long sociolinguistic interviews, an Ethnic Orientation Questionnaire (EOQ), and a picture naming task for 40 speakers of each of eight different heritage languages spoken in the GTA[32]. Each interview as well as the responses to the EOQ and picture naming task were digitally recorded making it possible to collect formant measurements for vowels uttered during any of these three sets of recordings for each speaker. The Cantonese interviews were conducted from 2009 to 2010 by Cantonese-speaking student researchers. To complement the recordings of Toronto speakers, a set of interviews of Homeland speakers from Hong Kong conducted in 2015 are also included. These interviews were conducted following the same procedures as the interviews conducted in Toronto.

In addition to the .wav file recordings, the HerLD Corpus also includes time-aligned transcriptions produced by native speakers (including HL speakers) using the program ELAN (Sloetjes and Wittenburg 2008). The Cantonese data was transcribed using the Jyutping Romanization System[33]. Instances of code-switching and code-mixing with English, which were both very common in the interviews, were transcribed using English orthography. In ambiguous cases, such as English lexical items pronounced using Cantonese phonology, the deciding criterion was whether or not the word appears in CantoDict[34] (an online crowdsourced dictionary of Cantonese using Jyutping). If the word appears in this dictionary with associated Chinese characters, the word is considered integrated enough into Cantonese to be a Cantonese word rather than a case of code-mixing with English. This study will only consider words transcribed in Jyutping and hence words that are recognized as integrated into Cantonese whether they are loans or part of inherited vocabulary.

---

[32] Aside from Cantonese, the other languages in the corpus include Faetar, Hungarian, Italian, Korean, Polish, Russian, and Ukrainian.

[33] Transcriptions with traditional Chinese characters are currently being added.

[34] http://www.cantonese.sheik.co.uk/scripts/wordlist.htm

The vowels that are examined in the current study come primarily from the sociolinguistic interviews. Vowel tokens also come from the EOQ and the Picture Naming Task if these recordings were available for a particular speaker and if the total length of the sociolinguistic interview for a speaker was less than an hour long. For interviews that were more than an hour long, tokens were included only up to the 60-minute point of the interview. The Picture Naming Task involved the use of a picture book that was shown to each participant. The participant was asked to provide a word in Cantonese to describe each picture of an isolated item and tell a story about each depicted scene.

The speakers analyzed can be divided into three groups for a grand total of 32 speakers: Homeland speakers (n = 8), GEN 1 Toronto speakers (n = 12), and GEN 2 Toronto speakers (n = 12). The Homeland speakers include speakers who were born in Hong Kong and have since lived continuously in Hong Kong. The Homeland speakers represent a group of speakers who have spent their entire lives in a place where Cantonese is the dominant language of everyday life. Knowledge of English is variable but does increase in apparent time primarily through schooling. GEN 1 includes those who 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. Since exposure to Toronto English would be during their adult years, any effect of English on Cantonese is expected to be minimal. Thus, the GEN 1 speakers are not expected to be different from the Homeland speakers in terms of how English could affect their Cantonese. GEN 2 speakers include those whose parents would qualify as GEN 1 speakers (even if those parents are not in the corpus). The HLVC project criteria included those who have arrived in the GTA before the age of six to be included. All but three of the GEN 2 speakers included in the present study were born in the GTA. Of the three speakers not born in the GTA, two arrived at the age of 2 while the third arrived at the age of 4.

Thus, all GEN 2 speakers had early exposure to both conversational Cantonese and conversational Toronto English. This makes the potential effect of Toronto English on their Cantonese phonology more likely than is the case with the other groups.

Table 19. Speakers Examined (N=32), with Year of Birth in Parenthesis

		1920s-1930s	1940s-1950s	1960s-1970s	1980s-1990s
HK (N=8)	M (N=3)			CXM52A(1963)	CXM27A(1988) CXM20A(1994)
	F (N=5)	CXF77A(1938)[35]		CXF49A(1965) CXF43A(1971)	CXF19A(1996) CXF16A(1998)
GEN 1 (N=12)	M (N=6)	C1M87A(1922)[36]	C1M61A(1949) C1M59A(1949) C1M52B(1956) C1M52A(1957)	C1M46A(1963)	
	F (N=6)	C1F83A(1926) C1F78A(1931)	C1F58A(1950) C1F54B(1954) C1F50B(1958) C1F50A(1959)		
GEN 2 (N=12)	M (N=6)			C2M44A(1965)	C2M27A(1983) C2M22A(1987) C2M21D(1987) C2M21C(1987) C2M21B(1987)
	F (N=6)			C2F41A(1969)[37]	C2F24A(1985) C2F22A(1986) C2F21C(1988) C2F21B(1988) C2F20A(1988)

[35] Born in Bao'an County, Guangdong (now part of Shenzhen), borders Hong Kong, which used to be part of Bao'an County before British colonization
[36] Born in Guangzhou
[37] Born in Hamilton, Ontario, not officially part of the GTA, but not far away

For each of the three groups, an attempt was made to select speakers representing as wide a range of ages as possible. Speakers in the GEN 1 group were born between 1922 and 1963 while speakers in the GEN 2 group were born between 1965 and 1988. The year of birth for the Homeland speakers spans from 1938 to 1994. This overlaps with the years represented by the GEN 1 and GEN 2 groups making it possible to address whether or not any differences observed between GEN 1 and GEN 2 are also changes in apparent time within the Homeland group.

Table 19 lists all of the speakers from the corpus that will be analyzed. Separate columns divide the list of speakers based on decade of birth. Each speaker in the corpus is identified by a code such as "C2F16A." The first character is a "C" indicating "Cantonese." The second character is a number indicating Generational Group. For Toronto speakers, this is a "1" for "GEN 1" or a "2" for "GEN 2." Homeland speakers are indicated with an "X". The third character is either an "M" for "male" or an "F" for "female". This is followed by the age of the speaker at the time of recording. The last character is a letter used to distinguish between different speakers in the corpus with the same demographic characteristics. For example, the speaker code C2F16D indicates a Cantonese-speaker that is second-generation, female, and 16 years of age. The "D" at the end indicates that there are three previously recorded speakers with these exact same demographic characteristics. The other three speakers would have the following speaker codes: C2F16A, C2F16B, and C2F16C. The birth year for each speaker is also included in parenthesis following each speaker code. This makes it easier to see the range of birth years represented since not all of the recordings were completed during the same year.

Other factors that may be important but that are not considered in the current study include socio-economic class background, neighborhood of residence, and languages spoken other than English and Cantonese. The GEN 1 and GEN 2 speakers recorded live throughout the

GTA.[38] Some speakers live within the city of Toronto. Others live in various ethnoburbs. Some speakers have also lived in more than one neighborhood within the GTA. The speakers are also involved in a variety of professional occupations. This includes social workers, engineers, architects, students, and receptionists among many others. There is also one speaker that is a Christian pastor. The Homeland group also includes a wide range of professions such as a retired factory work, a stay-at-home mom, a nurse, an IT project manager, a journalist, a flight attendant, and a student. With such a wide variety of professions represented and speakers living in many different neighborhoods within the GTA and in Hong Kong, the speakers examined do not appear to be a homogeneous group with respect to socio-economic class. This is exactly what is desired for the purpose of the current study since the main focus is on differences defined by generational group (distinguished by distinct language acquisition experiences) rather than on differences based on socio-economic class or neighborhood of residence.

Knowledge of languages other than English and Cantonese is another factor not considered. Information about other languages is incomplete and was not consistently recorded. Yet, as was discussed in Chapter, there are not any reasons to believe that knowledge of these other language would result in community-wide change among the specific group of Toronto speakers selected. Younger Hong Kong speakers, however, may show influence from Mandarin. The current study, however, has not been designed to test any hypotheses about how specific Mandarin vowels would affect Cantonese vowels.

[38] See map at http://projects.chass.utoronto.ca/ngn/HLVC/4_1_map.php

5.3.2 Ethnic Orientation Scores

At the end of each sociolinguistic interview, each participant was given an Ethnic Orientation Questionnaire (EOQ) consisting of a set of questions designed to address the extent to which each speaker is oriented to a specific ethnic identity. For the current study, the EOQ had two purposes. First of all, since speakers provided oral and digitally recorded responses to these questions, this made it possible to include vowel tokens uttered during EOQ responses to also be included in the analysis. This was particularly important for less talkative speakers who would have otherwise had a smaller number of analyzable vowel tokens. The second and more general purpose of the EOQ was to provide a systematic way of operationalizing individual ethnic identity in the analysis of vowel variation.

Table 20. EOQ Scores for each GEN 2 speaker

Speaker	EOQ Score (rounded to two decimal places)
C2F21C	0.68
C2M21D	0.69
C2F20A	0.71
C2F41A	0.75
C2M21B	0.75
C2F24A	0.77
C2M22A	0.78
C2M27A	0.86
C2F22A	0.88
C2F21B	1.00
C2M21C	1.10
C2M44A	1.30

The EOQ included a series of questions, each belonging to one of several categories related to ethnic orientation. These categories include ethnic identification, language, language choice, cultural heritage, parents, partner, homeland culture, and discrimination (see Nagy, Chociej, and Hoffman 2014 for more details). Not all individuals answered every question. One reason was because some questions such as questions about one's partner or children are not applicable to all participants. Although questions were open ended, responses to each question were coded with a numerical score of 0, 1, or 2. A score of "0" is for responses that indicate the highest level of assimilation into Canadian society, while a score of "2" is for the opposite end. A score of "1" is for responses that fall in the middle. For example, one question is "would you rather live in a Chinese neighborhood?" A definite "yes" response would be coded as a "2" while a definite "no" would be coded as a "0". Responses that show no strong preference would be coded as a "1".

Once all responses were coded, an average value was calculated for each individual speaker. This average value will henceforth be referred to as the "EOQ score". This value can be used as a continuous variable in modeling the extent to which ethnic orientation may be a factor favoring extreme articulations for the F1 or F2 of a particular vowel[39]. The EOQ scores for all of the GEN 2 speakers are shown in Table 20.

Individual speaker responses to some of these questions can also be considered in the analysis by addressing whether there is a relationship between specific responses and vowel pronunciations along the F1/F2 space. For instance, do speakers who orient more towards Canadian culture in a specific way lead in shifting a specific vowel in one direction or in innovating a phonologically conditioned split?

---

[39] Another approach is to divide speakers into low vs. high EOQ groups as in Hoffman and Walker (2010). I also tried this approach but it resulted in very similar results.

For the current study, the EOQ questions considered in statistical modeling include only those for which all GEN 2 speakers provided a response. There also must be a minimum of two different responses and a minimum of two speakers for each response. Questions that meet these requirements include questions A1, A2, A5, B1, B2, B5, C1, C4 and E2. The responses to these questions are summarized in the tables that follow.

Table 21. Do you think of yourself as Chinese, Canadian, or Chinese-Canadian **(**EOQ A1)

	Canadian (0)	Chinese-Canadian (1)	Chinese (2)
GEN 1		9 (75%)	**3 (25%)**
GEN 2	5 (42%)	**5 (42%)**	2 (17%)

Table 22. Are most of your friends Chinese? **(**EOQ A2)

	No (0)	Mixed (1)	Yes (2)
GEN 1		2 (17%)	**10 (83%)**
GEN 2	1 (8%)	**7 (58%)**	4 (33%)

Table 23. When you were growing up, were the kids in your school Chinese? **(**EOQ A5)

	No (0)	Mixed (1)	Yes (2)
GEN 1[40]			**11 (100%)**
GEN 2	3 (25%)	**5 (42%)**	4 (33%)

Table 24. Do you speak Cantonese? How well? **(**EOQ B1)

	No (0)	Speaks a little (1)	Speaks often (2)
GEN 1		1 (8%)	**11 (92%)**
GEN 2		**9 (75%)**	3 (25%)

[40] One GEN 1 speaker did not answer this question.

Table 25. Do you prefer to listen to the radio or watch TV in Cantonese or English? (EOQ B5)

	English (0)	Both (1)	Cantonese (2)
GEN 1 (N = 12)	2 (17%)	**8 (67%)**	2 (17%)
GEN 2 (N = 12)	**8 (67%)**	4 (33%)	

Table 26. What language does your family speak when you get together? (EOQ C1)

	English (0)	Both (1)	Cantonese (2)
GEN 1 (N = 12)	1 (8%)	5 (42%)	**6 (50%)**
GEN 2 (N = 12)	2 (17%)	**7 (58%)**	3 (25%)

Table 27. Do you speak to your parents in Cantonese or English? (EOQ C4)

	English (0)	Both (1)	Cantonese (2)
GEN 1 (N = 12)			**11 (100%)**
GEN 2 (N = 12)		**3 (25%)**	9 (75%)

Table 28. Do your parents speak Cantonese or English? (EOQ E2)

	English (0)	Both (1)	Cantonese (2)
GEN 1 (N = 12)[41]		1 (14%)	**6 (86%)**
GEN 2 (N = 12)	2 (17%)	**8 (67%)**	2 (17%)

Some of the questions that do not meet the established criteria are questions that had unanimously homogeneous responses from GEN 2 speakers. This includes the responses to EOQ questions B3, B4, and C2. The responses to each of these questions are summarized in Table 29, Table 30, and Table 31 respectively. For example, for the question "do you prefer to speak Cantonese or English?" as well as the question "what language do you speak with your friends",

[41] Data is missing from five GEN 1 speakers. This is probably because the answer to this question was too obvious for the interviewers that the interviewers did not bother to ask.

all except for one GEN 2 speaker said “English”. For the question, “do you prefer to read and write in Chinese or English?”, the distinction between GEN 1 and GEN 2 responses was completely categorical. Every GEN 2 speaker said “English” while GEN 1 speakers said either “Cantonese” or “both”. Since the responses to these questions do not make it possible to address how a subset of GEN 2 speakers may be innovative, these questions will not be considered in statistical modeling. The responses to these questions, however, do show how distinct language usage patterns categorically (or almost categorically) distinguish GEN 1 speakers from GEN 2 speakers.

Table 29. Do you prefer to speak Cantonese or English? (EOQ B3)

	English (0)	Both (1)	Cantonese (2)
GEN 1 (N = 12)		5 (42%)	**7 (58%)**
GEN 2 (N = 12)	**11 (92%)**	1 (8%)	

Table 30. Do you prefer to read and write in Chinese or English? (EOQ B4)

	English (0)	Both (1)	Chinese (2)
GEN 1 (N = 12)		**8 (67%)**	4 (33%)
GEN 2 (N = 12)	**12 (100%)**		

Table 31. What language do you speak with your friends? (EOQ C2)

	English (0)	Both (1)	Chinese (2)
GEN 1 (N = 11)[42]		3 (27%)	**8 (73%)**
GEN 2 (N = 12)	**11 (92%)**		1 (8%)

[42] One GEN 1 speaker is missing a response for this question.

5.3.3 Language Percentage and Word Count Scores

Although some of the EOQ questions discussed above relate to language use, the responses to these questions are all self-reported. Furthermore, as discussed above, the GEN 2 responses to some of these questions were completely (or almost) categorical. For these reasons, I also consider language use measures that can be operationalized as continuous variables. These measures are based on how speakers actually use their different languages in spontaneous speech. For instance, could it be the case that speakers who code-switch the most or who use English the most be the ones most likely to produce innovative variants?

Lyskawa et al. (2016) adopted one such measure in their study of Toronto Heritage Polish word-final devoicing. They computed a score by tallying the number of code switches in each interview and dividing the result by 60 to obtain a code-switching rate score. The results from this study showed that Heritage Polish speakers with higher code-switching rates had higher rates of word-final devoicing. These speakers devoiced in environments that do not favor devoicing in Homeland Polish. This suggests that "frequent code-switching provides the context in which these speakers' knowledge of Polish and English patterns converge" (Lyskawa et al. 2016:219).

In this dissertation, I consider three similar sets of measures that gauge the relative amount of Cantonese and English used by each speaker. These three measures will be referred to as the CAN % Score, the CAN WC Score, and the ENG WC Score. The advantage of these three measures is that they can be easily computed without the additional coding required for calculating the code-switching score used in Lyskawa et al (2016). Each score is based on the total number of Cantonese or English word tokens uttered. In tallying the number of Cantonese words, each individual syllable uttered was treated as a distinct word. This conforms to the common practice in Chinese linguistics of treating each syllable as a semantically meaningful phonological unit. The use of the Jyutping Romanization system in transcription made it easy to distinguish between Cantonese and English words. Since each Cantonese syllable must have a tone number, Cantonese words can easily be identified in spreadsheet formulas by the presence of these numbers while all other syllables can be easily identified as English.

The CAN % Score measures the percentage of the recorded speech sample that consists of Cantonese word tokens. This score was calculated by taking the total number of Cantonese word tokens and dividing it by the total number of all word tokens in the recordings for that speaker. The CAN WC Score is simply a count of the total number of distinct Cantonese words in the recordings for a specific speaker. It provides a rough measure of a speaker's Cantonese vocabulary size based on the number of unique vocabulary items uttered during the interview. The ENG WC Score is the same measure but for the number of unique English words uttered.

Some speakers occasionally uttered phrases in other languages such as Mandarin. In each case, however, the amount of speech in other languages was negligible and limited to metalinguistic commentary. English was consistently more common than any language other than Cantonese in the interviews. Languages other than English and Cantonese were, thus, excluded in the calculation of these three scores.

Table 32 shows the average CAN % and word count scores for all three groups analyzed. The Hong Kong group clearly has the highest CAN % score while the GEN 1 Toronto group is not far behind at 97.11%. The GEN 2 average CAN % score is almost 20 percentage points lower, but the majority of the GEN 2 recordings is still in Cantonese. Although the HK group had the highest CAN % score, it is the GEN 1 group that has a higher CAN WC score. In any case, both groups had speakers producing an average of more than 600 unique Cantonese words. For the GEN 2 group, the average was only 472 unique Cantonese words. This is still higher than the average ENG WC score for the GEN 2 group, which is also the highest across the three groups. Although there has been a lot of discussion about the influence of English in Hong Kong as mentioned in Chapter 4, the Hong Kong group averaged only 18 in the ENG WC score. The GEN 1 group was much higher at 163.

Table 32. Measures of language usage in interviews (averages)

	CAN %	CAN WC	ENG WC
GEN 1 Average	97.11	694	163
GEN 2 Average	77.89	472	403
HK Average	99.64	603	18

Table 33. Measures of language usage in interviews (range in values)

	CAN %		CAN WC		ENG WC	
	Low	*High*	*Low*	*High*	*Low*	*High*
GEN 1	91.85	99.51	530	960	18	430
GEN 2	29.36	98.59	308	715	100	1003
HK	99.21	99.94	499	781	4	59

Also helpful in gauging overall Cantonese usage levels is the range of scores for each group. This is shown in the boxplots that follow in Figure 4 (CAN % Scores), Figure 5 (CAN WC Scores), and Figure 6 (ENG WC Scores). Numerical values are shown in Table 33. A complete list of scores for all GEN 2 speakers is included in Table 70 in Appendix A. The Hong Kong group has a very narrow range of scores for CAN %. The lowest value was 99.21% while the highest value was 99.94%. The GEN 1 group had a larger range of variation for this metric.

Still, all speakers had at least 91.85% of their speech sample in Cantonese. The GEN 2 group had the largest range of variation not only for the CAN % score but also for the ENG WC Score. In contrast, the HK group used very little English in the recordings.

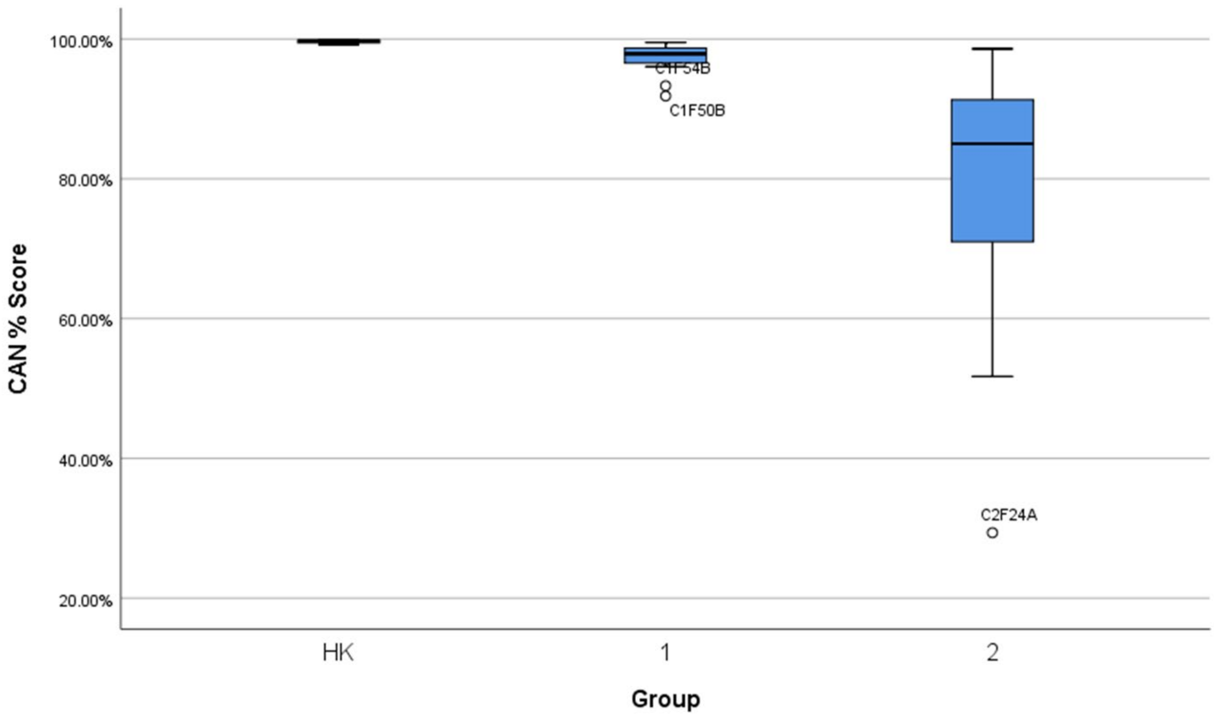

Figure 4. Range of CAN % Scores across three groups

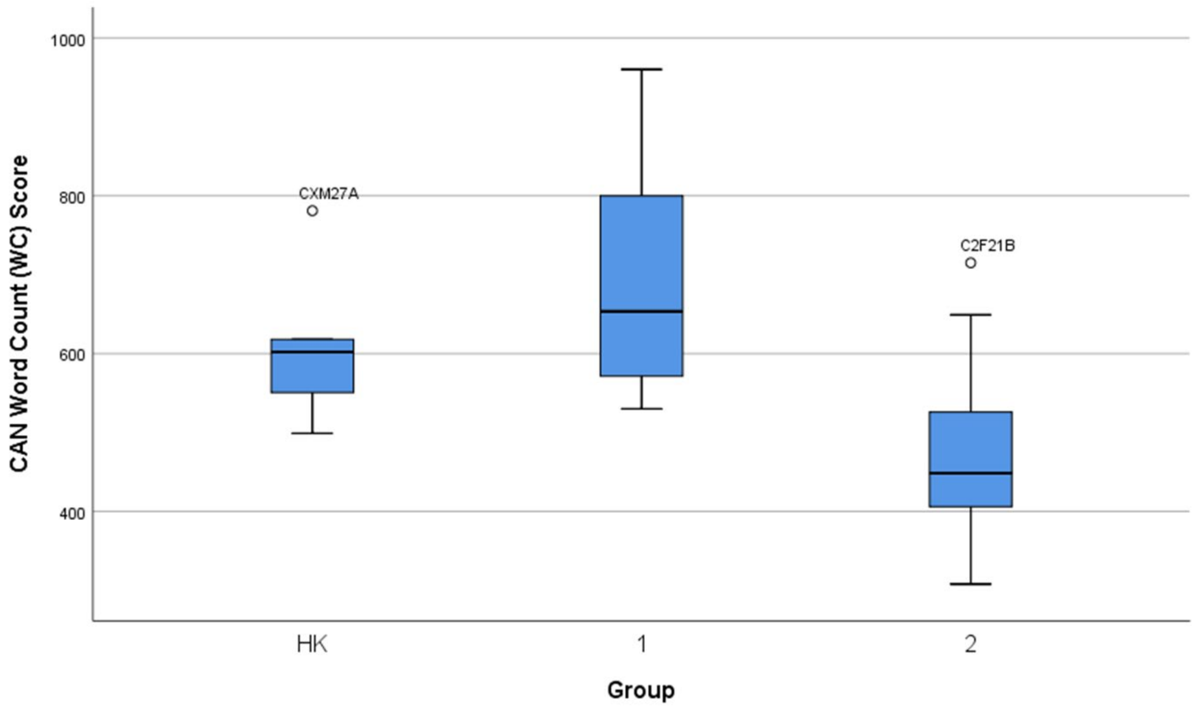

Figure 5. Range of CAN WC Scores across three groups

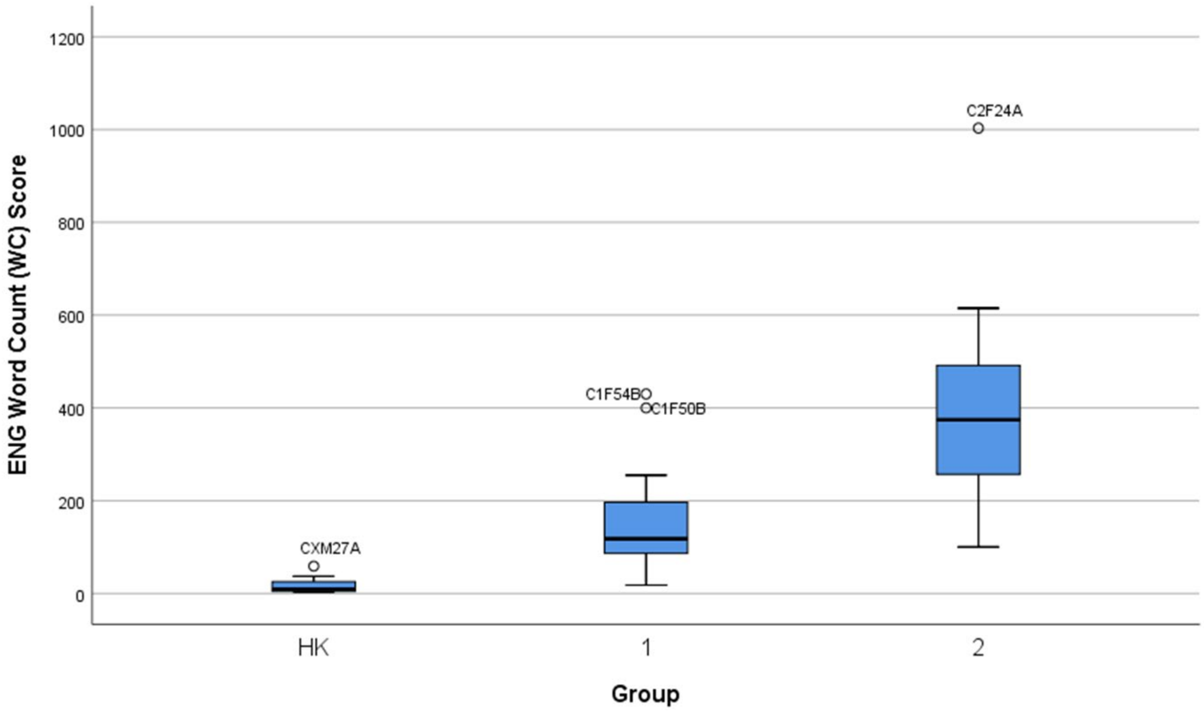

Figure 6. Range of ENG WC Scores across three groups

5.4 METHODOLOGY AND PROCEDURES

The process for identifying usable vowel tokens, extracting formant measurements, and coding each token involved eight steps. These steps included transcription preparation (Section 5.4.1), Prosodylab-Aligner (Section 5.4.2), manual review of textgrids (Section 5.4.3), formant extraction by script (Section 5.4.4), output review (Section 5.4.5), vowel plot visualization (Section 5.4.6), normalization (Section 5.4.7), and preparing spreadsheet for Rbrul (Section 5.4.8).

5.4.1 Transcript Preparation

The process began with the ELAN transcripts that have been previously transcribed and proofread by various native speakers (including HL speakers). These transcripts, with the .eaf file extension, are part of the HerLD corpus. The first step was to export the .eaf transcript to a Praat textgrid file. A Praat script was then run on the textgrid file to split the entire transcript into smaller chunks. Each of these chunks corresponded to a single annotation in ELAN. For each chunk, the output included both a .wav file of the annotation and a matching .txt file. The .txt file contained the transcribed text for the annotation. In general, each annotation corresponded to a single sentence or utterance. To prepare for the next step, the .txt files were then converted to .lab format, which is a text format readable by the program Prosodylab-Aligner (Gorman, Howell, and Wagner 2011).

5.4.2 Prosodylab-Aligner (Automated)

The second step was to perform speech to segment alignment of the data using the program Prosodylab-Aligner. Unlike FAVE[43] (Rosenfelder et al. 2011), which has become increasingly popular in sociolinguistics research on English dialects, Prosodylab-Aligner can be customized to work on any language. It has three requirements. First, Prosodylab-Aligner needs a minimum of one-hour of recorded speech (.wav) so that it can be trained to work on any arbitrary set of data. Second, each .wav file must have a corresponding .lab file, which is essentially a type of .txt file which includes the transcription of the utterance in a format readable by Prosodylab-

[43] "Forced Alignment and Vowel Extraction"

Aligner. Finally, in order to interpret the mapping between orthography and audio features, Prosody-lab requires a dictionary written following the format of the CMU Pronouncing Dictionary[44]. A customized dictionary for Cantonese was created (Cui 2014; Cui et al. 2014). This dictionary was combined with the existing CMU Pronouncing Dictionary for English. This combined file made it possible for Prosodylab to deal with the use of English since code-switching to English in the interviews was common.

The output of running Prosodylab was a set of Praat textgrids for each corresponding .wav/.lab file pair. These textgrids had all of the phonemes automatically labeled (based on the Jyutping transcription) and automatically aligned to the waveform and spectrogram. One study showed that Prosodylab-Aligner has an accuracy rate of about 80% for Cantonese data (Peters and Tse 2016). Thus, some manual correction was needed.

5.4.3 Manual Review of Alignment and Formant Tracker Accuracy

The third step of the process was to identify specific tokens for analysis. Tokens selected were indicated in the word tier of each textgrid file with a pair of slashes enclosing the chosen syllable. If the syllable is part of a compound word, the rest of the compound word was also added. In identifying specific tokens, only syllables with one of the 11 Cantonese monophthongs were chosen. Syllables with onset glides ([j] and [w]) or labio-velar co-articulated stops ([k^w] and [kw^h]) were excluded. Vowels in open syllable context immediately followed by a syllable with an onset glide without an intermediate pause (ex: [hɔ2 ji5], 'able to') were also excluded. Since the focus of this study is on change in terms of vowel categories rather than in terms of lexical

[44] http://www.speech.cs.cmu.edu/cgi-bin/cmudict

diffusion, a maximum of 10 tokens for each word per speaker were selected. A separate spreadsheet was used to keep track of the total number of tokens for each word.

For each token identified, the phoneme boundaries of the vowel were adjusted if needed in the Prosodylab-Aligner-generated textgrids. The Praat formant tracker was also set to appropriate values as shown in Table 34. If the formant tracker was off target for a portion of the duration of a vowel, the boundaries were adjusted to include only the accurate portions of the vowel. Tokens were excluded in cases involving undershoot, overlapping speech, laughter, singing, too much background noise, unusually rapid speech, or other problems that make reliable formant measurements difficult or impossible.

Table 34. Praat Formant Settings

	Male	Female
Maximum Formant	5000 Hz	5500 Hz
# of Formants	5	5
Window Length	0.025 seconds	0.025 seconds

5.4.4 Formant Extraction Script

After all of the textgrids were manually reviewed, a Praat script was run on all of the usable .textgrid/.wav file pairs to automatically extract the values of the first two formants for all of the selected monophthongs in the audio. The output file included a list of all of the tokens extracted along with file location information, vowel category, and F1 and F2 measurements for each token.

5.4.5 Output Review

The fifth step was to review the output generated by the formant extraction script for errors. To facilitate the identification of errors, the file was sorted by vowel category and then by increasing or decreasing F1 and F2 values. This made it possible to review the entire list of tokens for errors. The sound files containing the tokens with the three lowest and three highest F1 and F2 values were carefully reviewed to ensure that the values extracted were accurate. Errors were subsequently corrected if possible. Otherwise, these tokens were removed from the output spreadsheet and hence excluded from analysis.

One example of a common problem was syllable fusion, a phonological process in Cantonese connected speech that involves the fusing together of a di-syllabic compound word or phrase into a single syllable (Wong 2006). The fused syllable often contains a vowel and tonal contour that are intermediate between the two vowels in the full form of the compound word or phrase.

Another common problem was formant tracker errors. An example is shown in Figure 7 below. In this image, the entire duration of the vowel /a/ (represented in Jyutping as AA) in the word <saam1> (‘three’) is shown. For most of the duration of this vowel, the F1 aligns with the lowest dark band while the F2 aligns with a higher dark band. Near the midpoint, however, are three formants rather than two. This part of the image is circled. The second formant is an extra formant that was miscalculated. A case like this results in the F2 being calculated as the F3 and an F2 measurement that is too low. If possible, vowel boundaries were adjusted to include only accurate portions. Otherwise, such tokens were excluded from analysis.

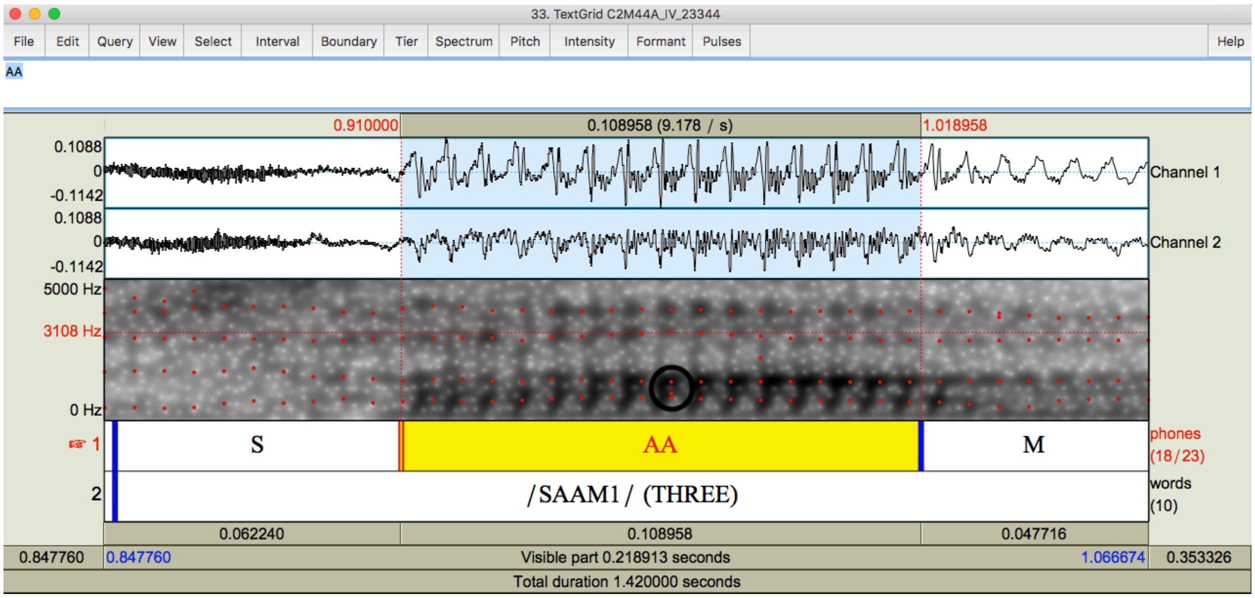

Figure 7. Example of extra formant (shown in circle)

5.4.6 Vowel Plot Visualization

The sixth step was to create vowel plots for each speaker based on raw formant values using the vowel normalization suite NORM (Thomas and Kendall 2007). Each vowel plot included ellipses for each vowel category that indicate the mean F1/F2 ±1 standard deviation. The question that needed to be addressed in this step was whether or not each speaker produces seven distinct tense vowels and four distinct lax vowels based on F1/F2 values. Vowel categories were considered distinct if they showed lack of overlap in the ellipses. This was important because some vowel normalization techniques are based on assumptions of a shared set of phonological contrasts. The answer to this question determined the normalization procedure used in the next step.

5.4.7 Normalization

The results from the vowel plots showed that all speakers produce 11 acoustically distinct vowels. The Lobanov technique in the NORM suite (Thomas and Kendall 2007) was chosen. It is a vowel extrinsic and speaker intrinsic normalization method. (Thomas 2011:165) discusses the advantages and disadvantages of several different techniques. The disadvantages of vowel-intrinsic techniques include dependence on either F0 or F3 and distortion of the vowel space. Since this study uses spontaneous speech samples, many of which were recorded in suboptimal conditions, F3 measurements may not be as reliable as those from recordings completed in more controlled settings. The distortion of the vowel space is also a problem for the current study since one of the goals is to address the implications of the results of this study for variationist sociolinguistics theory. For this reason, it would be preferable to have normalized data comparable to what has been used in most variationist studies. Thus, a vowel-extrinsic technique is preferable. Thomas (2011) identifies two disadvantages of vowel-extrinsic techniques. The first is that they work optimally when the entire vowel system is included. This is exactly the case for the current study so this is not a concern. The second disadvantage is that the results may be impaired if different vowel systems are normalized together. As was shown in the previous step, all speakers have the same set of 11 contrasts. Thus, this disadvantage is not relevant either.

The other consideration in choosing a normalization technique is speaker-intrinsic vs. speaker-extrinsic. There is only one speaker-extrinsic technique available in the NORM suite. It is the one used for the *Atlas of North American English* (Labov, Ash, and Boberg 2006). The disadvantage of this technique is that it works optimally with at least 345 speakers. The number of speakers analyzed for the current study is much smaller. Thus, after considering all of the advantages and disadvantages I have chosen a vowel-extrinsic and speaker-intrinsic method. Among the vowel-extrinsic and speaker-intrinsic techniques offered through the NORM suite, Adank, Smits, and Van Hout (2004) showed that the Lobanov technique worked the best. The Lobanov technique was also the technique used in previous studies of Toronto Heritage Cantonese (Tse 2016a; 2016b). Thus, I used the Lobanov technique again for the current study so that the results can be more directly comparable to these previous studies.

Having all of the data normalized together made it possible to compare speakers from different generational groups as well as speakers from two different places. The output of NORM was a new tab delimited text file with normalized values for the first two formants for each token along with transcriptions and speaker identifiers.

5.4.8 Preparing Spreadsheet for Rbrul

Finally, the last processing step was to merge the NORM output into a spreadsheet with all of the independent variables examined. Since token information was linked to speaker code and Jyutping orthography, Excel formulas were used to quickly code independent variables related to the word and to the speaker. For instance, with a token with speaker code C2F16B, formulas were used to extract the variants (or levels) for Generation (in this case GEN 2), Sex (in this case Female), and age (in this case 16). In addition, each speaker code was linked to EOQ and language usage scores. For a token of the syllable "GAA1" formulas extracted information for the variable "onset" (in this case /g/), "vowel" (in this case AA or /a/ in IPA), "coda" (in this case #, meaning open syllable), and "Tone" (in this case 1). After all of the values for each variable and for each token were entered, the spreadsheet was saved as a tab delimited text file and uploaded to the program Rbrul (Johnson 2009) for mixed effects modeling.

5.5 VARIABLES FOR STATISTICAL MODELING

To reiterate, the research questions are as follows:

(Q1) Is there evidence for contact-induced inter-generational vowel shifting in native vocabulary?

(Q2) Is there evidence for contact-induced vowel mergers or vowel splits in native vocabulary? Four specific parts of the Cantonese vowel space are addressed:

(Q2a) Is there evidence for a merger between /y/ and /u/?

(Q2b) Is there evidence for an increasing acoustic split between /i/ and /ɪ/?

(Q2c) Is there evidence for an allophonic split in /ɛ/?

(Q2d) Is there evidence for an allophonic split in /ɔ/?

(Q3) To what extent can demographic, ethnic orientation, or language use factors (Sex, Age, EOQ Score, specific EOQ responses, CAN % Score, CAN Vocabulary Score, ENG vocabulary score) account for the propagation of the specific shifts, mergers, and splits observed in the data?

For each vowel category, two separate analyses were run with either F1 or F2 as the dependent variable. I will henceforth refer to this dependent variable as a “formant/vowel pair.” Unless otherwise noted, all statistical analyses included both random effects and mixed effects as independent variables. The random effects for each model include both “Speaker” and “Word”.

Treating these variables as random effects is a way of controlling for the possible effect that individual speakers or individual words may have on the overall results. The fixed effects included depended on the specific research question. The variables that will be considered, along with a brief description of each variable, are listed below.

<u>Speaker:</u> This refers to individual speaker. There are hence 32 different values. "Speaker" will be included as a random effect in most of the models that will be run.

<u>Syllable:</u> This refers to each distinct syllable.

<u>Word:</u> This includes syllables that are part of multisyllabic words as well as monosyllabic words. This will be included as a random effect in most of the models that will be run.

<u>Generational Group:</u> Possible values include GEN 1, GEN 2, and HK.

<u>Onset:</u> This corresponds to the consonant that immediately precedes the vowel if present. Words without consonant onsets will be coded as having a zero-onset and represented as "#". This is one of the phonological conditioning factors that will be considered in addressing Q2 and Q3.

<u>Coda:</u> This corresponds to the consonant that immediately follows the vowel if present.

<u>Tone:</u> This refers to the tone category of the word in which a vowel token occurs. This will be coded as a number (1, 2, 3, 4, 5, or 6) corresponding to the tone category included in the Jyutping transcription of the word.

Based on results of "onset", "coda", and "tone", it may be possible that specific natural class groupings of consonants or tone categories may be more informative. If this is the case, post-hoc tests could include variables such as "velar coda" instead of simply "coda". Similarly, level vs contour tone could also be shown to be a meaningful distinction in post-hoc models.

Sex: This indicates whether a token comes from a speaker identified as male ("M") or female ("F").

Age: This indicates the age of the speaker at the time of recording, which would have been 2009 or 2010 for the Toronto speakers and 2015 for the Homeland speakers.

EOQ Score: This variable was explained in Section 5.3.3.

Individual EOQ Responses: In addition to an overall EOQ for each speaker, responses to individual questions are also available for each speaker. The EOQ questions that are most relevant to language were described in Section 5.3.3. These responses can also be used to address whether or not the innovative speakers are the ones with particular responses to specific EOQ questions. Each of these EOQ variables is modeled as a categorical variable. Due to collinearity, no more than one EOQ variable was included in each model.

A1: Do you think of yourself as Chinese, Canadian, or Chinese-Canadian?

A2: Are most of your friends Chinese?

A5: When you were growing up, were the kids in your school Chinese?

B1: Do you speak Cantonese? How well?

B5: Do you prefer to listen to the radio or watch TV in Cantonese or English?

C1: What language does your family speak when you get together?

C4: Do you speak to your parents in Cantonese or English?

E2: Do your parents speak Cantonese or English?

Finally, the last set of variables are variables for the CAN % Score, CAN WC Score, and ENG WC Score discussed in Section 5.3.3. As is the case for the individual EOQ responses, no more than one of these variables were included in each model due to collinearity. EOQ responses, CAN % Score, and Word Count Scores were considered only in analyses of variation within the GEN 2 group.

6.0 RESULTS

I begin this chapter with descriptive statistics related to the results (Section 6.1) followed by a discussion of the overall vowel space for the three speaker groups examined (Section 6.2). The first research question (Q1) is addressed in Section 6.3: Is there evidence for inter-generational vowel shifting? Results addressing the second research question (Q2) are presented in Section 6.4: Is there evidence for vowel mergers or vowel splits influenced by contact with Toronto English? The specific hypotheses formulated in Chapter 5 were (Q2a) merger of /y/ and /u/, (Q2b) increasing split of [i] vs. [ɪ], (Q2c) split in /ɛ/, and (Q2d) split in /ɔ/. Results addressing the third research question (Q3) are presented in Section 6.5: To what extent can demographic, ethnic orientation, or language use factors (Sex, Age, EOQ Score, specific EOQ responses, CAN % score, CAN vocabulary score, ENG vocabulary score) account for the specific shifts, mergers, and splits observed in the data? This chapter concludes with a results summary in Section 6.6.

6.1 DESCRIPTIVE STATISTICS OF TOKENS ANALYZED

The results presented in this chapter are based on an analysis of the complete set of usable tokens processed following the procedures discussed in Chapter 5. Figure 8 shows the total number of tokens included for each vowel and for each of the three groups. Although the GEN 1 and GEN

2 groups each included the same number of speakers (12 each for a total of 24), the GEN 1 group had more usable tokens than the GEN 2 group. The HK group included the smallest number of speakers (N=8) as well as the smallest number of usable tokens (N = 7,491). The grand total of usable vowel tokens from all groups is 33,179.

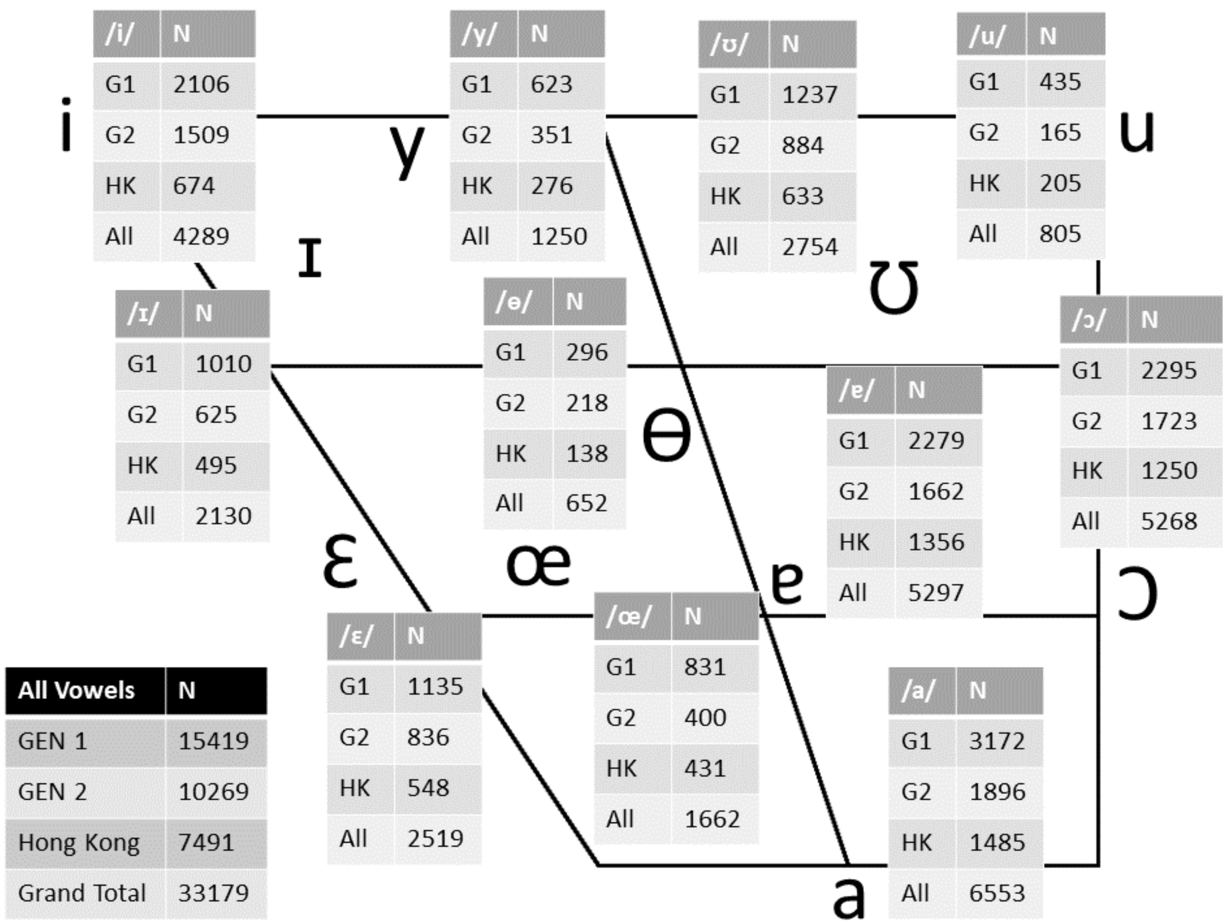

Figure 8. Total number of vowel tokens for each vowel category and for each group

Table 35 shows the percentage of total tokens for each group that is represented by each vowel category. Vowel categories are listed based on the percentage ranking for the GEN 1 group. The ranking of the most common vowels is similar across all three groups. The only difference between GEN 1 and GEN 2 is the relative ranking of the two least frequent vowels, /u/ and /ɵ/, while the only difference between GEN 1 and the Homeland group is the relative ranking of /ɔ/ and /ɐ/, both quite small discrepancies. These vowels are indicated in bold. In both cases, the relative ranking is switched around. With similar rankings across all three groups, the speech samples analyzed across all three speaker groups appear to be comparable.

Table 35. Percentage of total tokens for each vowel

Vowel	GEN 1	GEN 2	HK
/a/	20.57%	18.46%	19.82%
/ɔ/	**14.88%**	16.78%	**16.69%**
/ɐ/	**14.78%**	16.18%	**18.10%**
/i/	13.66%	14.69%	9.00%
/ʊ/	8.02%	8.61%	8.45%
/ɛ/	7.36%	8.14%	7.32%
/ɪ/	6.55%	6.09%	6.61%
/œ/	5.39%	3.90%	5.75%
/y/	4.04%	3.42%	3.68%
/u/	**2.82%**	**1.61%**	2.74%
/ɵ/	**1.92%**	**2.12%**	1.84%

6.2 OVERALL VOWEL SPACE

F1 and F2 means for each vowel category across the three groups are included in the vowel plots that follow with showing the tense vowels Figure 9 and Figure 10 showing the lax vowels. Each ellipse represents one standard deviation from the mean F1 and F2 of each vowel for each group. The mean F1/F2 is represented with a red dot for the GEN 1 group, a blue empty square for the GEN 2 group, and a green triangle for the Homeland (HK) group. Of the 33,179 total tokens included, 22,346 are for the tense vowels while 10,833 are for the lax vowels. Most of these 11 vowels are acoustically distinct in F1/F2 across all three groups. The two notable exceptions involve round vowels overlapping with unrounded vowels. For example, /i/ and /y/ show overlap for the GEN 1 and GEN 2 groups. Similarly, /ɐ/ and /ɵ/ also overlap. Since /y/ and /ɵ/ are round and /i/ and /ɐ/ are unrounded, the overlap in F1/F2 values for these vowels does not indicate merger.

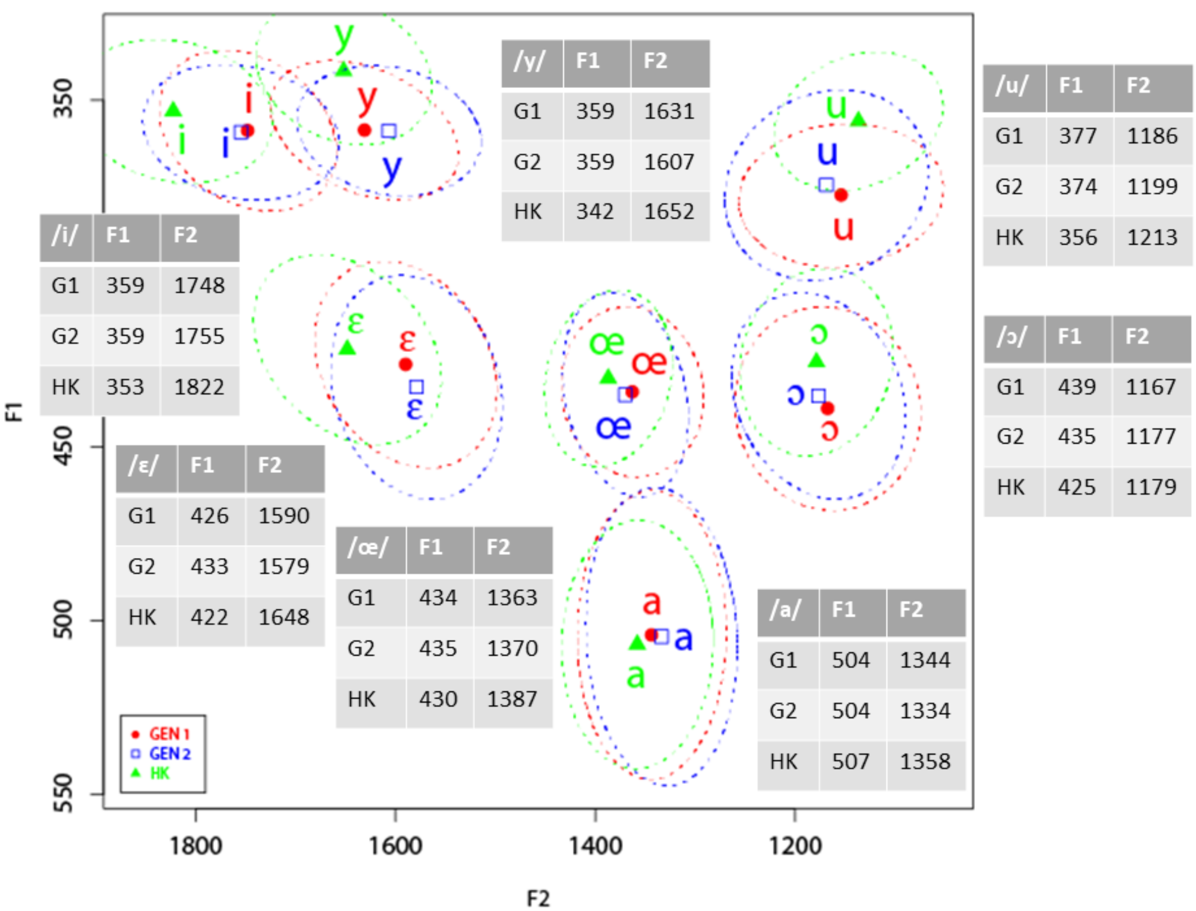

Figure 9. Tense vowels, 32 speakers, n=22,346. Ellipses indicate mean F1/F2 ± 1 SD

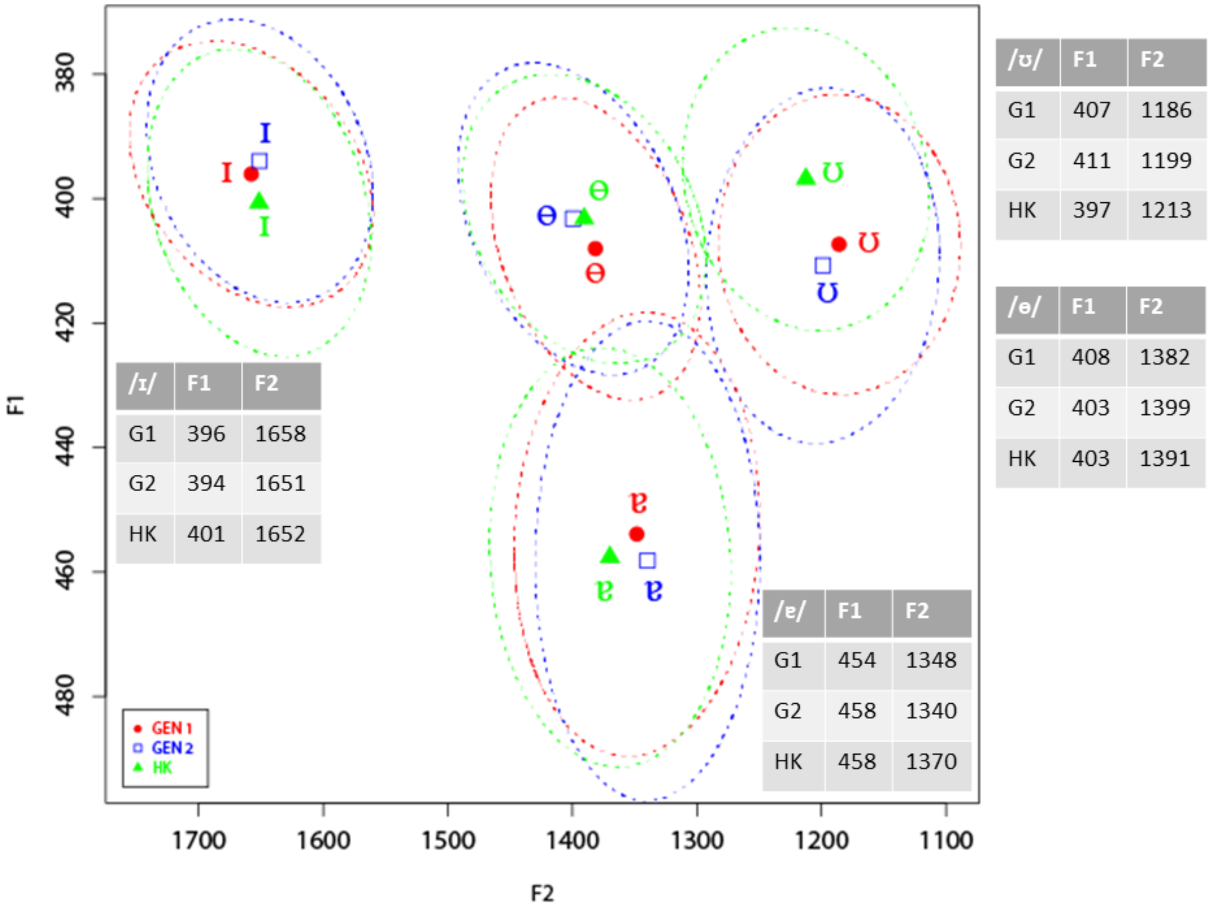

Figure 10. Lax vowels, 32 speakers, N=10,833. Ellipses indicate mean F1/F2 ± 1 SD (rounded to the nearest Hz)

The relative similarity between GEN 1 and GEN 2 in contrast to the HK group is immediately visible in these plots. In many cases, the GEN 1 and GEN 2 ellipses are closer together than either of them is to the HK ellipses. This is especially the case for /i/, /y/, /u/, and /ɛ/. These four vowels appear to be more peripheral[45] for the HK group than they are for either

[45] The peripherality of three of these vowels (/i/, /y/, and /ɛ/) was also observed in Hong Kong in a study comparing Hong Kong and Guangzhou speakers (Lee 1983).

GEN 1 or GEN 2. They are raised, fronted, or both raised and fronted in comparison to their counterpart vowels in the Toronto groups.

The mean formant values shown in Figure 9 and Figure 10 also show similarity between the two Toronto groups. GEN 1 and GEN 2 have the exact same mean F1 values (rounded to the nearest whole number) for /a/. This is also the case for /i/ and /y/. The GEN 2 and Homeland group also share some of the exact same F1 means but only for two vowels (/ɐ/ and /ɵ) as opposed to three. In general, the inter-group differences between GEN 1 and GEN 2 are smaller than they are for differences between the HK and Toronto groups. For instance, while the difference between the HK and GEN 2 group for the F2 of /i/ is 67 Hertz, the biggest GEN 1 vs. GEN 2 difference is only about 24 Hertz (for /y/). In fact, as I will show in the next section, /y/ is the only vowel that shows a significant inter-generational difference in formant values. The results in Section 6.3 also confirm that there is more vowel shifting in the HK group.

6.3 EVIDENCE FOR VOWEL SHIFTS? (Q1)

Q1 was addressed by running a set of mixed effects models that included either F1 or F2 as the dependent variable, “speaker” and “word” as random effects, and “group” (GEN 1 or GEN 2) as a fixed effect. Separate models were run for the F1 and F2 of each vowel category. Thus, with two formants for each of 11 vowels, a total of 22 different models were run. Results from these 22 different models are summarized in Table 36.

The only model showing statistical significance was the model for the F2 of /y/. Detailed information from this model is shown in Given substantial acoustic similarity between the two groups and no overlap in the age range included in each group (20-44 for GEN 2, 46-87 for GEN 1), one might ask if the retraction follows a shift initiated among GEN 1 speakers. In other words, could the retraction of /y/ be an internally motivated change that started among GEN 1 speakers rather than a contact-induced change initiated by GEN 2 speakers. Similarly, could there be other internally motivated changes that are not evident from models that include generational group as a categorical fixed effect?

Table 37. According to this model, the GEN 1 group has a tendency of producing higher F2 (mean of 1634 Hz) than the GEN 2 group (mean of 1608 Hz). This means that the GEN 2 group produces significantly more retracted variants of /y/ than the GEN 1 group. Thus, since only one vowel shows an inter-generational difference, we do not have evidence for a chain shift. Since only one of these 22 models came out significant with an alpha value set at 0.05, one may suspect a false positive. In Section 6.4.1, I present metalinguistic commentary about the pronunciation of this vowel along with evidence of a merger between /y/ and /u/. This suggests that GEN 2 retraction of /y/ is a change worthy of further investigation.

Given substantial acoustic similarity between the two groups and no overlap in the age range included in each group (20-44 for GEN 2, 46-87 for GEN 1), one might ask if the retraction follows a shift initiated among GEN 1 speakers. In other words, could the retraction of /y/ be an internally motivated change that started among GEN 1 speakers rather than a contact-induced change initiated by GEN 2 speakers. Similarly, could there be other internally motivated changes that are not evident from models that include generational group as a categorical fixed effect?

Table 36. Is "group" a significant predictor for each vowel/formant pair?

	Separate models for each vowel/formant pair For each model Random Effects: "Speaker" and "Word" Fixed Effect: "Group" (GEN 1 vs. GEN 2)	
	Is "group" significant for F1 as dependent variable?	Is "group" significant for F2 as dependent variable?
/y/	n.s.	* ($p < 0.05$, r^2 [fixed] = 0.05)
/ɐ/	n.s.	n.s.
/a/	n.s.	n.s.
/ɛ/	n.s.	n.s.
/ɵ/	n.s.	n.s.
/i/	n.s.	n.s.
/ɪ/	n.s.	n.s.
/ɔ/	n.s.	n.s.
/œ/	n.s.	n.s.
/u/	n.s.	n.s.
/ʊ/	n.s.	n.s.

Table 37. Mixed effects model for the F2 of /y/ for Toronto groups only

Random Effects ($r^2 = 0.368$) Speaker, Word			
Fixed Effect ($r^2 = 0.05$) Group ($p = 0.0389$)*	Coefficient	Tokens	F2 Mean (Hz)
GEN 1	24	623	1631
GEN 2	-24	351	1607
r^2 [total] = 0.418			

To address these two questions and to build a more solid case for contact-induced change, 22 different models were run with the same factors as for the ones run to produce the results shown in Table 36 but with "age" as a continuous fixed effect instead of "group" as a categorical fixed effect. The results of these models are shown in Table 38. The model for the F2 of /y/ did not come out significant, which suggests that the retraction of /y/ is not an internally motivated change initiated by GEN 1 speakers. The only model that came out significant was the model for the F2 of /i/. This model (details shown in Table 39) shows an inverse relationship between age and the F2 of /i/, which means the younger the speaker, the more likely they are to produce /i/ with higher F2 (more fronting). Thus, unlike the retraction of /y/, it appears that the fronting of /i/ is an internally-motivated change initiated by GEN 1 speakers.

Table 38. Is "age" (continuous) a significant predictor of variation for each vowel/formant pair?

	Separate models for each vowel/formant pair Only GEN 1 and GEN 2 data included in these models <u>For each model:</u> Random Effects: "Speaker" and "Word" Fixed Effect: "Age" (continuous)	
	Significant Predictor for F1?	Significant Predictor for F2?
/y/	n.s.	n.s.
/ɐ/	n.s.	n.s.
/a/	n.s.	n.s.
/ɛ/	n.s.	n.s.
/ɵ/	n.s.	n.s.
/i/	n.s.	*, $p = 0.0256$, r^2 [fixed] = 0.046
/ɪ/	n.s.	n.s.
/ɔ/	n.s.	n.s.
/œ/	n.s.	n.s.
/u/	n.s.	n.s.
/ʊ/	n.s.	n.s.

Table 39. Mixed effects model for the F2 of /i/ with GEN 1 and GEN 2 data included

Random Effects ($r^2 = 0.304$) Speaker, Syllable		
Fixed Effect ($r^2 = 0.049$) Age ($p = 0.0227$)* Age range included: 20-87		Coefficient
Continuous	+1	-1.069
r^2 [total] = 0.353		

To address whether Homeland speakers show evidence of the same changes in apparent time, a set of models like the ones run to produce the results shown in Table 38 were run on the Hong Kong data. The results from these models are summarized in Table 40. The model of "age" as a dependent variable for the F2 of /y/ did not come out significant. Since the retraction of /y/ is not an apparent time change in Hong Kong, a contact-induced change account for the retraction of /y/ in Toronto seems more promising. The fronting of /i/, identified above as an apparent time change in Toronto, is also an apparent time change in Hong Kong. As shown in Table 41, there is an inverse relationship between age and F2 values which means that younger speakers produce /i/ with higher F2 (more fronted articulations). The other models with "age" as a continuous fixed effect that came out significant were the models for the F1 of /ɪ/ (Table 42), the F1 of /ɔ/ (Table 43), F2 of /ɔ/ (Table 44), and the F1 of /ʊ/ (Table 45). All of them show an inverse relationship between age and formant values. This means that younger speakers produce vowels with higher F1 (articulatorily lowered) and higher F2 (fronted). Thus, lowering of /ɪ/ (Table 42), lowering of /ɔ/ (Table 43), fronting of /ɔ/ (Table 44), and lowering of /ʊ/ (Table 45) are all apparent time changes in the HK group. These directions of movement are also indicated in Table 40.

Again, it must be pointed out that running 22 models with an alpha of 0.05 raises the likelihood of either a Type 1 or Type 2 error. Although, this is true, it should also be noted that the changes in the HK group are consistent with the Principles of Vowel Chain Shifts. For instance, Principle I states that "tense vowels rise along a peripheral track" (Labov 1994:176). The triangular shape of the vowel space means that vowel raising co-occurs with vowel fronting. This triangle is described in Labov (1994:177) and illustrated for changes in Cantonese in Figure 11. The fronting of /i/, a tense vowel, would thus be consistent with Principle I. Principle II states that lax vowels lower and that is exactly the direction of movement shown by the two high lax vowels, /ɪ/ and /ʊ/, Finally, Principle III states that back vowels move to the front. The fronting of /ɔ/ observed would be consistent with Principle III. The simultaneous lowering movement coincides with a downward movement along that side of the triangular vowel space. While a closer investigation to confirm the development of each of these vowel shifts in Hong Kong would certainly be a worthwhile project, the main focus of this dissertation is on the possibility of contact-induced change in Toronto. What is important about the Hong Kong analysis is that the vowels showing evidence for shifting in apparent time are completely different from the one and only shift in Toronto that appears to be contact-induced.

Table 40. HK results for "age" (continuous) as significant predictor of variation

	Separate models for each vowel/formant pair Only Hong Kong data included in these models <u>For each model</u> Random Effects: "Speaker" and "Word" Fixed Effect: "Age" (continuous), Age range: 16-77		
	Significant Predictor for F1?	Significant Predictor for F2?	Direction of Change (if applicable)
/y/	n.s.	n.s.	--
/ɐ/	n.s.	n.s.	--
/a/	n.s.	n.s.	--
/ɛ/	n.s.	n.s.	--
/ɵ/	n.s.	n.s.	--
/i/	**n.s.**	*	**fronting**
/ɪ/	*	**n.s.**	**lowering**
/ɔ/	**	**	**Fronting/lowering**
/œ/	n.s.	n.s.	--
/u/	n.s.	n.s.	--
/ʊ/	*	**n.s.**	**lowering**

Table 41. Mixed effects model for the F2 of /i/

Random Effects ($r^2 = 0.126$) Speaker, Word		
Fixed Effect ($r^2 = 0.035$) Age ($p = 0.0268$)* Age range: 16-77		Coefficient (in Hertz)
Continuous	+1	-0.953
r^2 [total] = 0.161		

Table 42. Mixed effects model for the F1 of /ɪ/

Random Effects ($r^2 = 0.23$) Speaker, Word		
Fixed Effect ($r^2 = 0.054$) Age ($p = 0.0138$)* Age range: 16-77		Coefficient (in Hertz)
Continuous	+1	-0.31
r^2 [total] = 0.284		

Table 43. Mixed effects model for the F1 of /ɔ/

Random Effects ($r^2 = 0.381$) Speaker, Word		
Fixed Effect ($r^2 = 0.044$) Age ($p = 0.00868$)** Age range: 16-77		Coefficient (in Hertz)
Continuous	+1	-0.298
r^2 [total] = 0.425		

Table 44. Mixed effects model for the F2 of /ɔ/

Random Effects ($r^2 = 0.234$) Speaker, Word		
Fixed Effect ($r^2 = 0.017$) Age ($p = 0.00219$)** Age range: 16-77		Coefficient (in Hertz)
Continuous	+1	-0.511
r^2 [total] = 0.251		

Table 45. Mixed effects model for the F1 of /ʊ/

Random Effects (r^2 = 0.218) Speaker, Word		
Fixed Effect (r^2 = 0.019) Age (p = 0.0259)* Age range: 16-77		Coefficient (in Hertz)
Continuous	+1	-0.171
r^2 [total] = 0.237		

To conclude this section, the results show evidence for only one inter-generational vowel shift among Toronto speakers: the retraction of /y/. A model of "age" as a continuous fixed effect showed that the retraction of /y/ is not likely an internally motivated change initiated by GEN 1 speakers. Rather, it seems more likely to be a contact-induced change influenced by Toronto English. The lack of the same change in the Hong Kong data further supports a contact-induced change account. The only other change identified in the Toronto data was the fronting of /i/, which was also a change identified in the Hong Kong data. Thus, the fronting of /i/ seems to be a change already in progress among GEN 1 speakers that has been further advanced by younger Toronto speakers. The results also show more innovative vowel shifting in apparent time in the HK group than in the GEN 2 group. Other changes observed in the HK data that are absent in the Toronto data include the lowering of the lax vowels /ɪ/ and /ʊ/ as well as the lowering and fronting of /ɔ/. A graphical summary of all Toronto and Hong Kong vowel shifts reported is presented in Figure 11.

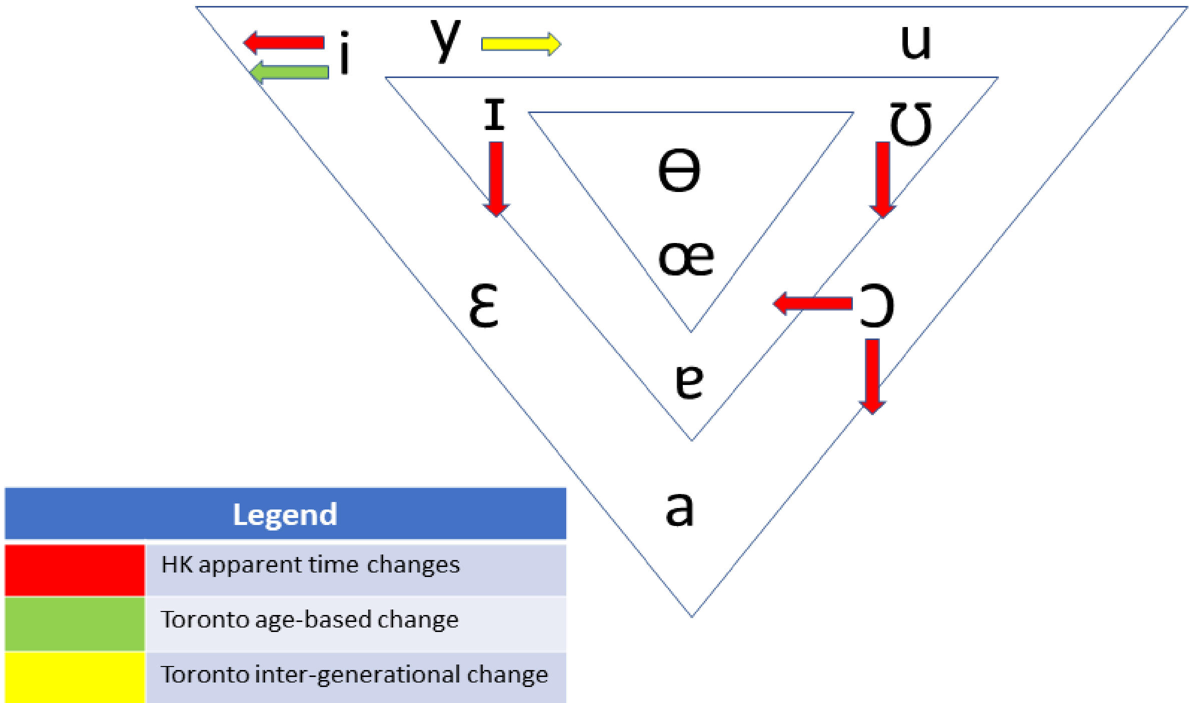

Figure 11. Summary of changes identified in data

6.4 MERGERS OR SPLITS? (Q2)

In this section, I present results addressing each of the four possible phonological changes formulated in Chapter 5. They are as follows:

a) The merger of /y/ and /u/

b) Increasing split of [i] vs. [ɪ]

c) A phonologically conditioned split in /ɛ/

d) A phonologically conditioned split in /ɔ/

6.4.1 Is there a merger of /y/ and /u/ (Q2a)?

The first hypothesized change I discuss is the possible merger of /y/ and /u/. Influence from Toronto English could mean the loss of a contrast between these two vowels since Toronto English has only one high round vowel. The retraction of /y/ and lack of change in /u/ reported in Section 6.3 suggests a change towards merger of these two vowel classes. The results I present below show that overall, GEN 2 speakers as a group have not merged /y/ and /u/. A few speakers, however, have brought the distribution of the two vowel classes close enough together to show evidence of a merger in progress (at least in production).

Table 46. Mixed effects model of /y/ vs. /u/ for GEN 2

Random Effects (r^2 = 0.043) Speaker, Word			
Fixed Effect (r^2 = 0.81) Vowel (p = 1.05e-55)***	Coefficient	Tokens	F2 Mean (Hz)
/y/	216	351	1607
/u/	-216	165	1169
r^2 [total] = 0.853			

In Table 46 are results from a mixed effects model that includes data from all /y/ and /u/ tokens from the GEN 2 group. The dependent variable is F2. The random effects include “speaker” and “word” while “vowel” is the only fixed effect. The model shows that “vowel” is a significant predictor of F2 (p*** < 0.001) with a huge r^2 [fixed] of 0.81. This suggests lack of merger. On the individual speaker level, all individual speaker mean F2 values for /y/ are above 1500 Hz while all mean F2 values for /u/ are below 1300 Hz (see Appendix A for detailed listing of mean F1/F2 for each individual speaker). Individual speaker means are illustrated in Figure

12.Although there is no overall merger for the GEN 2 group, a close examination of vowel token distributions reveals evidence for a merger for a few speakers. This appears to be related to the inter-generational retraction in /y/ reported in Section 6.3. To identify the specific speakers who have brought the distributions of the two vowel classes closest together, Pillai Scores (discussed in detail in Section 5.2.2) were calculated for each speaker. In Figure 13 are boxplots showing the range of Pillai Scores of the difference between /y/ and /u/ for each speaker group. As discussed in Section 5.2.2, Pillai Scores are based on a continuous scale ranging from 0 (lack of difference) to 1 (lack of similarity). All GEN 1 speakers had a Pillai score of 0.798 or above. The GEN 2 group, however, has the greatest range of Pillai Scores. This includes having the four lowest Pillai scores. A one-way ANOVA was run with Pillai scores as the dependent variable and group as the independent factor to determine if the inter-group differences in the range of Pillai scores is significantly different. LSD post-hoc tests showed that the only significant difference is that between the GEN 2 and the HK group. Most HK speakers have moved in the opposite direction of GEN 2 speakers by increasing the acoustic difference between these two vowels. The lack of speakers with lower Pillai scores in Hong Kong makes a contact-induced change explanation for a merger among a few GEN 2 speakers seem likely.

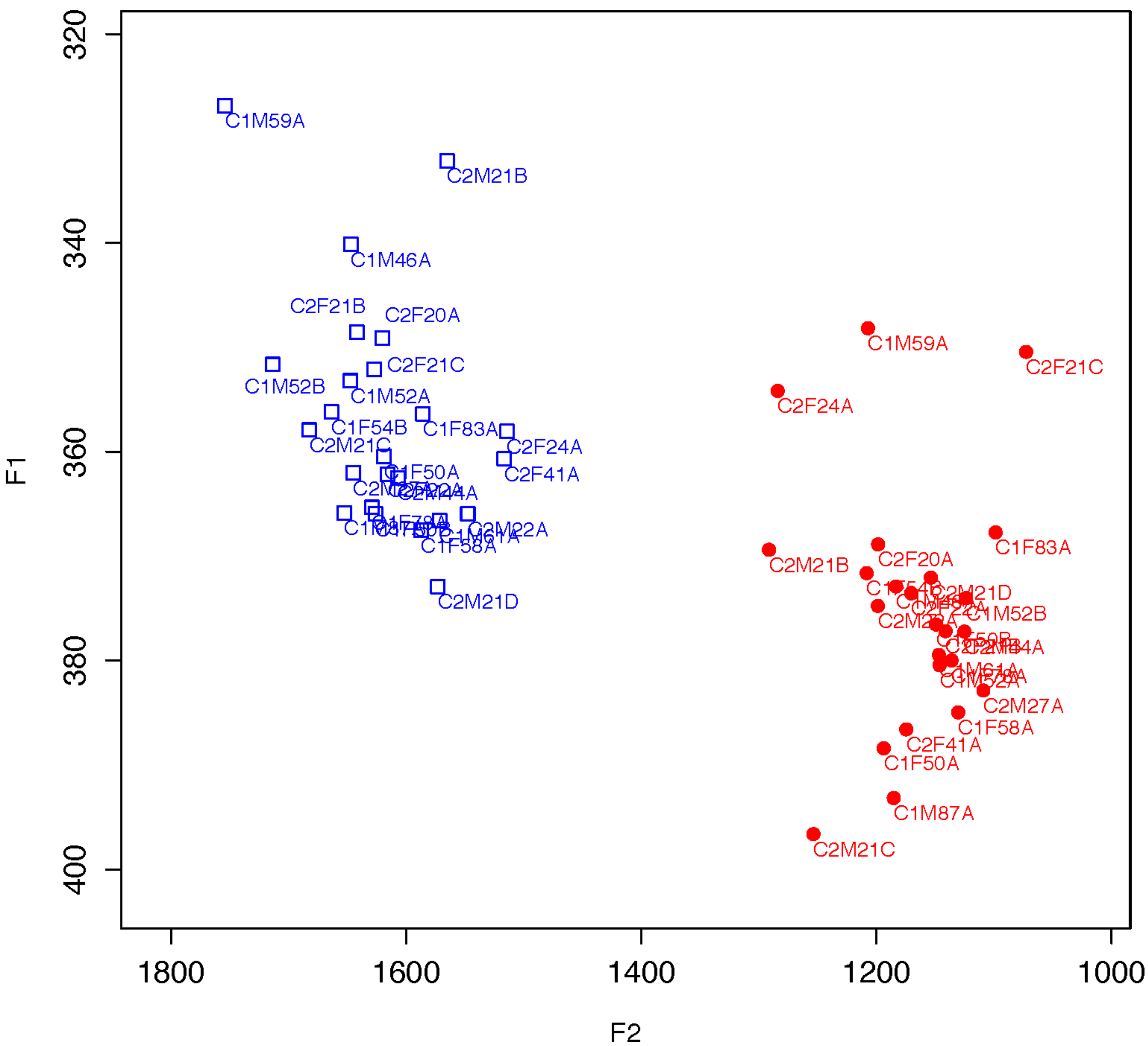

Figure 12. F1 and F2 means plotted for each Toronto speaker, /y/ in squares, /u/ in dots

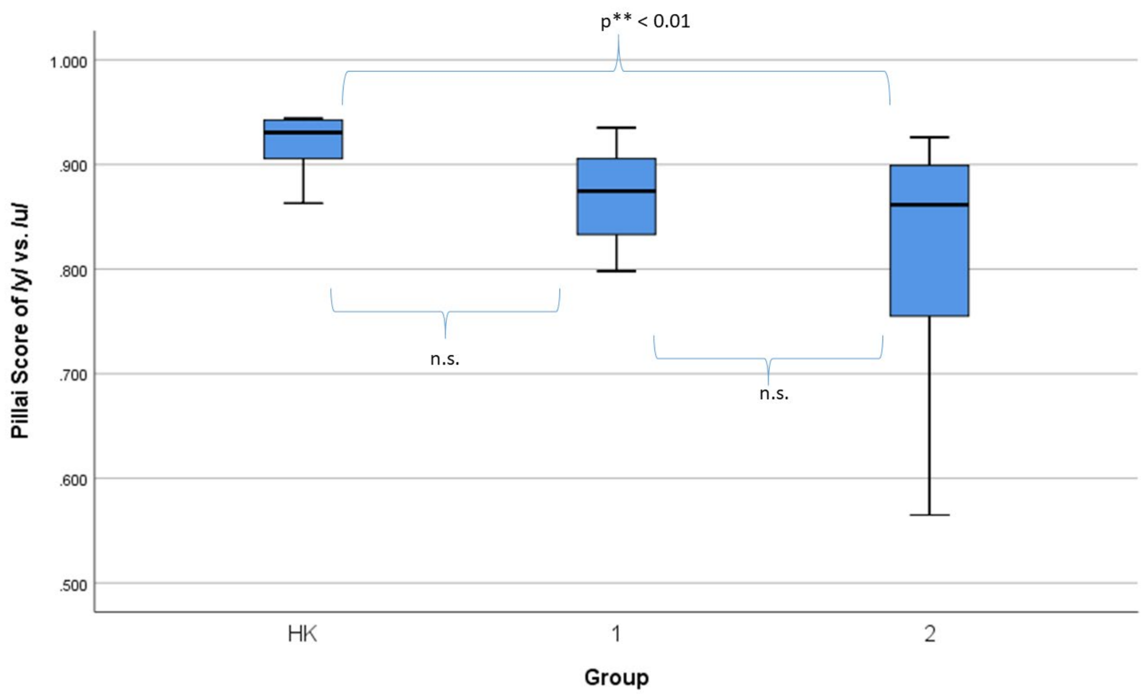

Figure 13. Range of Pillai Scores (/y/ vs. /u/) across three groups

Ideally, minimal pairs would be examined to determine whether a speaker has a merger. The current data set, however, lacks minimal pair tokens. Part of the problem is the near complementary distribution relationship of the two vowels. With coronal onset context, only /y/ is possible while in labial onset context, words can only have /u/. Velar onset context is the only context in which the two vowels contrast. An example of a minimal pair would be [kun35] <gun2> 'building' vs. [kyn35] <gyun2> 'roll'. The phonological distribution of the two vowels in Cantonese, thus, limits the number of possible minimal pairs. In the current set of data, minimal pair production data is available for only a few GEN 1 speakers.

Even without minimal pair data, the vowel plots of the speakers with the lowest Pillai Scores show partially overlapping vowel distributions. In Figure 14 below is the plot of all /y/

and /u/ tokens for C2M22A, the speaker with the lowest Pillai Score. At 0.565, this score places C2M22A near the middle of the continuum of being completely merged and being completely distinct. Most tokens of /y/ (represented as YU in the Jyutping Romanization system) are above 1400 Hz while most tokens of /u/ (represented as U in Jyutping) are under 1400 Hz. Between 1200 Hz and 1400 Hz is a cluster of three tokens of the word [cyun4] ('entire') circled in red. Also circled in red are the tokens of the vowel /u/ that have the highest F2 from this speaker. This includes one token of the word [bun1] ('to move') and one token of the word [fu3] ('pants or trousers'). Both of these /u/ tokens are more fronted than the cluster of [cyn4] tokens circled in red. Normally, we would expect /u/ to be more retracted than /y/, but that is not the case for the tokens circled in red. Thus, for this speaker, we can conclude that there is some overlap between the two vowel classes. This is evidence for a merger in production. In contrast, Figure 15 shows the plot of all /y/ and /u/ tokens from C2M44A, the speaker with the fourth highest Pillai Score (0.897) in the GEN 2 group. For C2M44A, there is zero overlap in the F2 range of the two vowel classes.

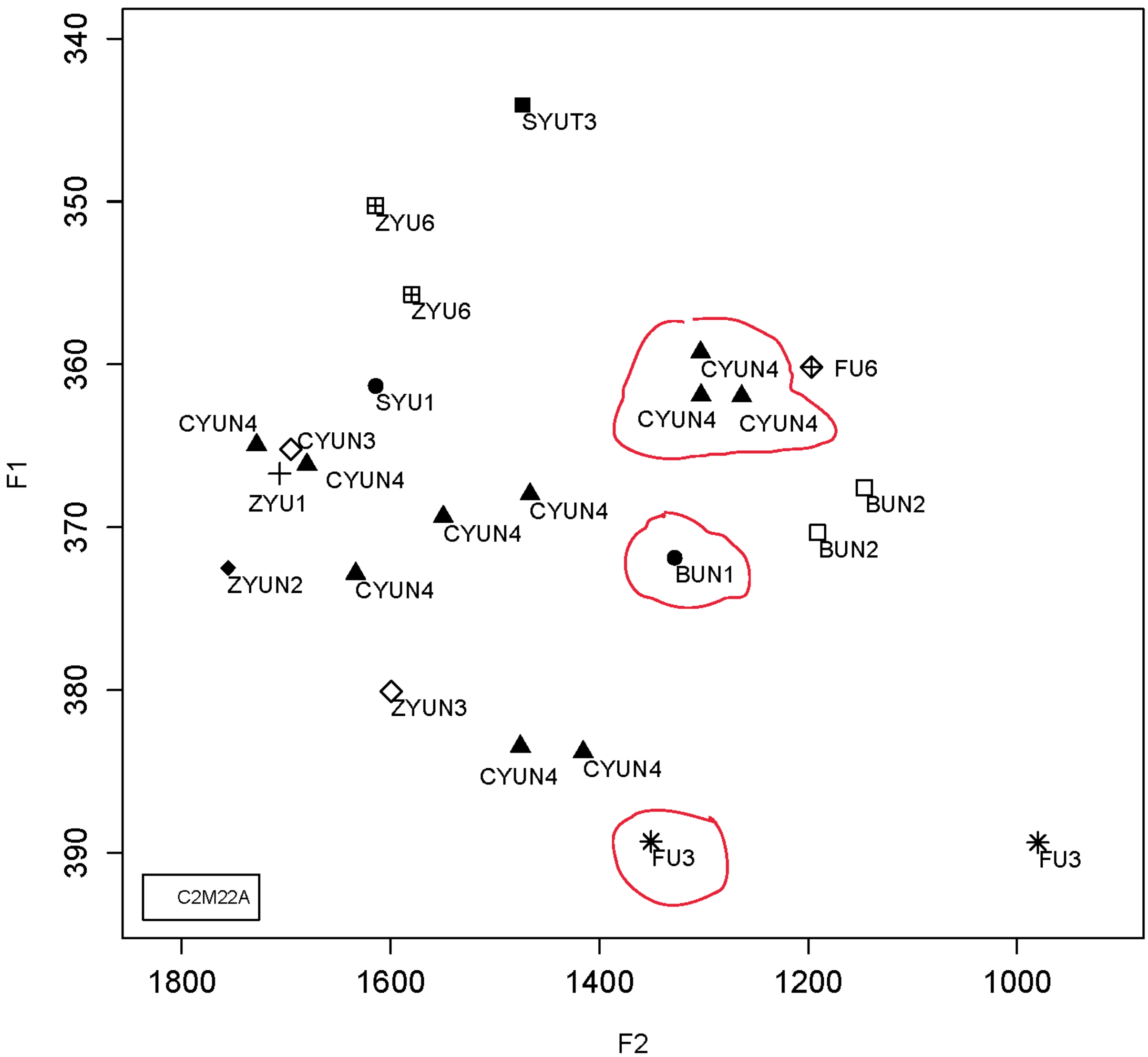

Figure 14. Plot of /y/ vs. /u/ for C2M22A (Pillai Score = 0.565), Lobanov normalized Hz values

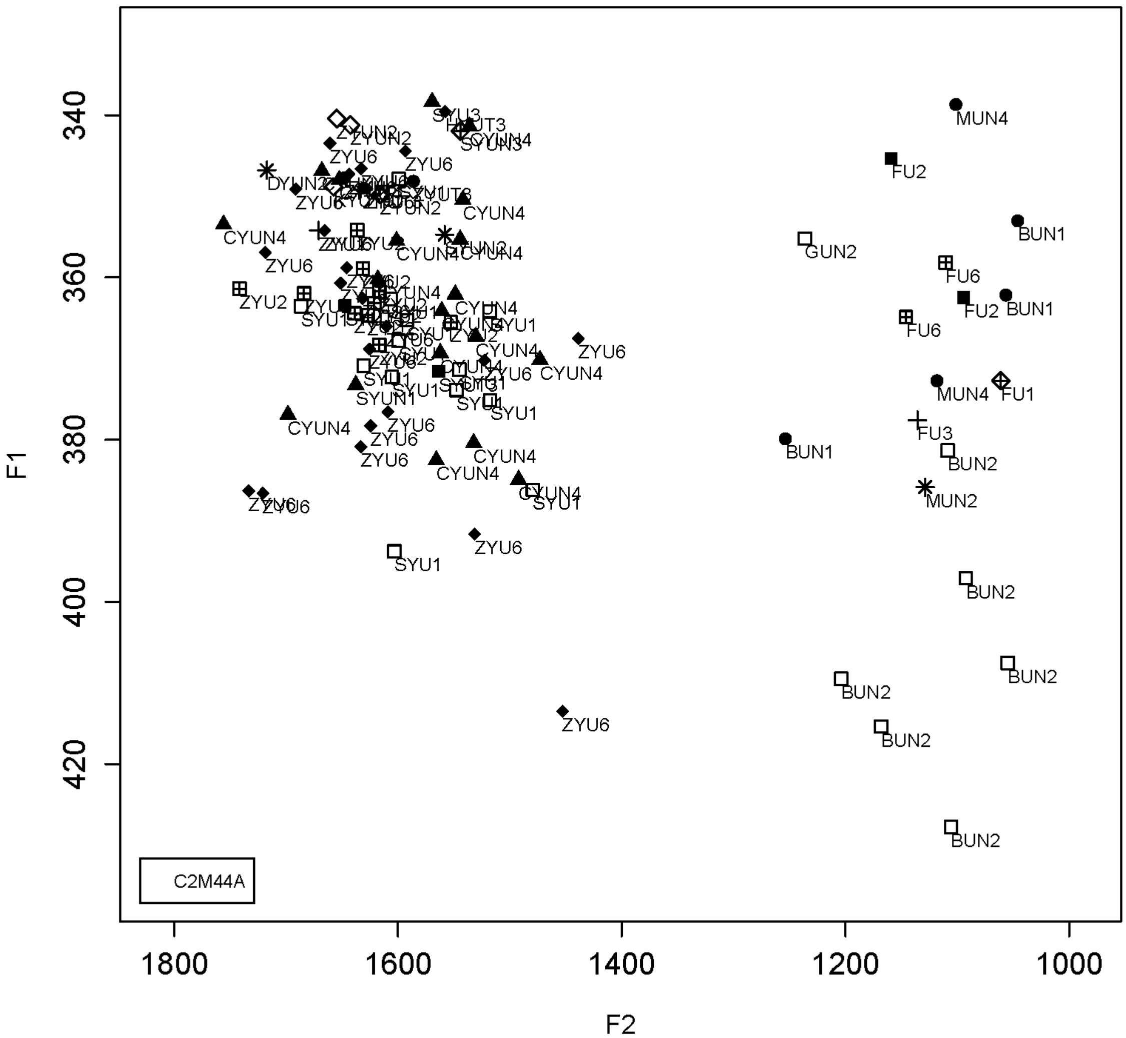

Figure 15. Plot of /y/ vs. /u/ for C2M44A (Pillai Score: 0.897), Lobanov normalized Hz values

In addition to production data, metalinguistic commentary from the speaker with the second lowest Pillai Score shows evidence of a merger in perception. This is shown in the excerpt below, which comes from the transcript for the picture naming task recording for C2F24A. For reference, the excerpt is immediately followed by a glossary of all Cantonese words included in the transcript. Tokens that appear in **bold** in the transcript are labeled in red on the vowel plot shown in Figure 16. The [ ] indicate overlapping speech, while the < > are used to indicate Cantonese vocabulary. Most of this excerpt is in English. In fact, C2F24A has the lowest CAN % Score and the highest ENG WC Score of any speaker analyzed. The vowel plot includes all /y/ and /u/ tokens measured from C2F24A. The excerpt begins with C2F24A identifying pictures of "ice cream" (<syut3 gou1>) and "fish" (<jyu2>)[46]. Both Cantonese words contain /y/[47]. C2F24A pronounces the /y/ in <syut3 gou1> with a Lobanov normalized F2 of 1507 Hz. Although this is not far from her average F2 for /y/, it is still 100 Hz lower than the GEN 2 mean and 124 Hz lower than the GEN 1 mean. Her retraction of /y/ became very salient to the interviewer. This elicited some discussion about her unusual pronunciation of certain words.

1. C2F24A: uh, <**syut3 gou1** … jyu2 … beng2> … what's so funny?
2. Interviewer: [(LAUGH)   ]
3. C2F24A: [Did I say it wrong?]
4. Interviewer: Your pronunciation
5. C2F24A: [What?]
6. Interviewer: [I can't] say <jyu2> (IMITATING C2F24A)
7. C2F24A: Oh I said it right?
8. Interviewer: No, wait say it again.
9. C2F24A: <jyu2>
10. Interviewer: OK.
11. C2F24A: "No, people say I say things weird [like] <**dau6 fu6**> or like **<zyu1>**
12. Interviewer: [yeah, it's]
13. C2F24A: they all [say] I say it wrong!"
14. Interviewer: [yeah]
15. Interviewer: <zyu1> and <jyu2, dau6 fu6> is right
16. C2F24A: I said <**dau6 fu6**> right?
17. Interviewer: Yeah, <zyu1> and <jyu2>, I think you said it wrong.
18. C2F24A: **<zyu1>**

---

[46] As mentioined in Chapter 5, words with initial glides were excluded from the current study. The formant measurements for tokens of <jyu2> were, thus, not included in the analysis. The metalinguistic discussion in this excerpt, however, makes it clear that the /y/ retraction also affects words in glide onset context.

[47] It was completely unintentional that the task included two consecutive words with Cantonese /y/ since the same task was designed to be used for multiple languages that are part of the HLVC Project.

19. Interviewer: [yeah! (LAUGH)]
20. C2F24A: [yeah! (LAUGH)]

Cantonese Glossary: <syut3 gou1> (“ice cream”), <jyu2> (“fish”), <beng2> (“cookie”), <dau6 fu6> (“tofu”), <zyu1> (“pig”), <zyu6> (aspect marker, shown in vowel plot), <dai3 fu3> (“underwear”, shown in vowel plot)

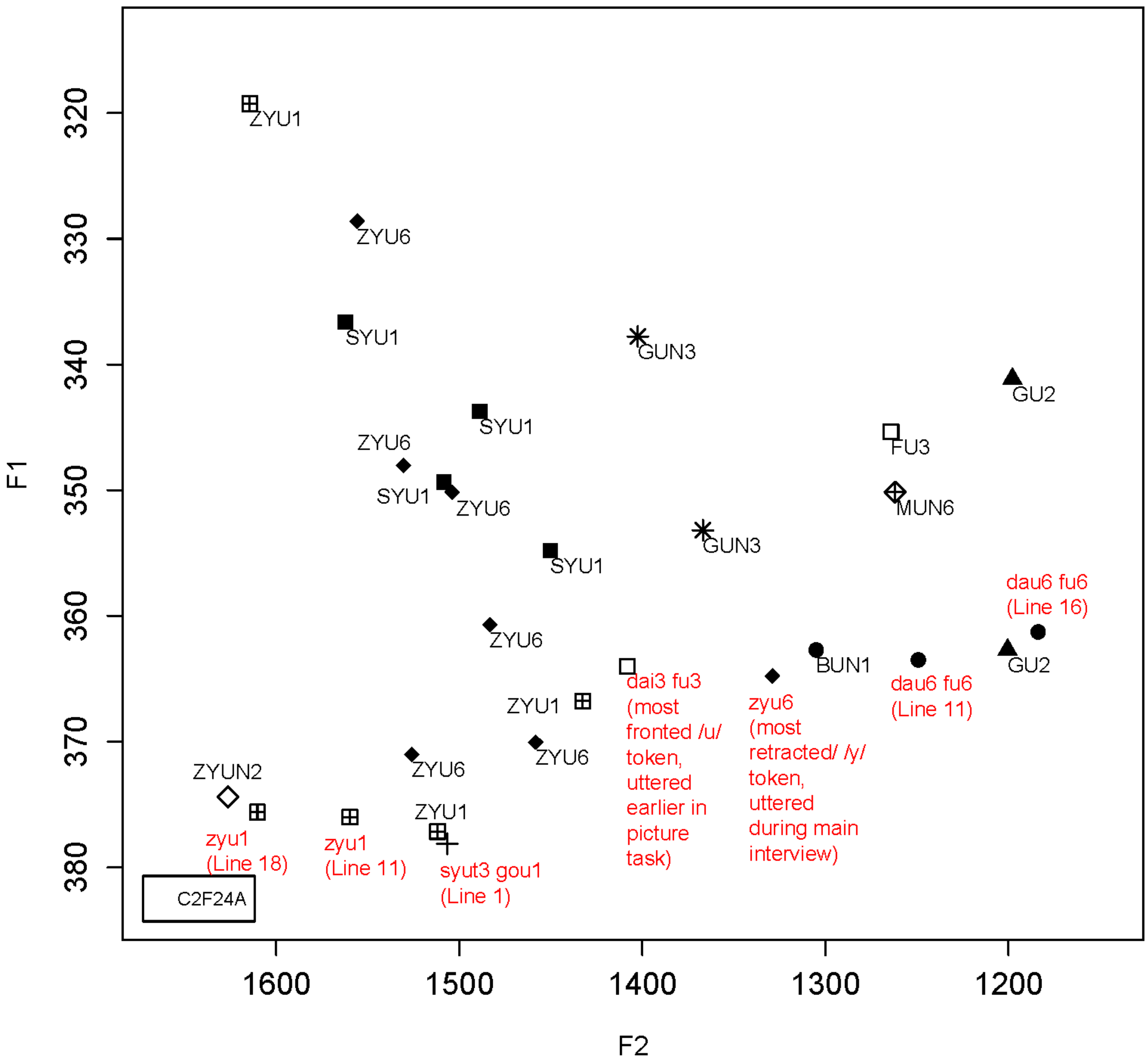

Figure 16. Plot of /y/ vs. /u/ for C2F24A (2nd lowest Pillai Score) with line numbers from transcript

What is particularly striking about this conversation is how the innovative pronunciation of one of the vowels involved with a possible merger (/y/) elicited examples of words with the two contrasting vowels (<dau6 fu6> and <zyu1>). It is clear that the interviewer is not the only person who has noticed C2F24A's innovative pronunciation of these two vowels. C2F24A also has the lowest F2 mean for /y/ of any of the 32 speakers analyzed. As shown in Figure 16, the most retracted token of /y/ (in the word <zyu6>) is more retracted than the most fronted token of /u/ (in the word <dai6 fu3>). There are also two tokens of the word <gun3> ('tin can or container') that are between the most retracted token of /y/ and the most fronted token of /u/. Thus, for C2F24A, there is partial overlap of the two vowel classes and evidence for a merger. <gun3> is also the only word uttered by C2F24A that could potentially form a minimal pair. Yet, it is pronounced with an unusually high F2 for /u/[48]. Does this mean C2F24A would pronounce the word <gyun3> with a similar F2 or with a higher F2 to maintain a contrast? If she pronounces it with the same F2 range as <gun3>, she would have a merger rather than simply an incipient merger.

On the other hand, C2F24A also shows evidence of contrast maintenance, though not in a phonetic context in which minimal pairs are possible. C2F24A's most retracted token of /u/ was produced in Line 16 when she asks if she said the word [dau6 fu6] correctly. One of her most fronted tokens of /y/ was produced shortly after in Line 18. Once the interviewer called attention to her pronunciation, C2F24A may have become more careful (perhaps unconsciously) in

[48] When I listened to the part of the recording in which these two tokens of <gun3> were uttered, I experienced perceptual confusion because of the extreme /u/ fronting. The /u/ impressionistically seemed too fronted for /u/ but at the same time too retracted for /y/. If it wasn't for the context of the conversation, I would have guessed that she said <gyun3 gyun3> ('coupon') or <gyun2 gyun2> ('rolls'). Although the latter has a different tone, it is also a more common word and would have also made sense in the context of the sentence, but less so in the context of the conversation.

enhancing the contrast between these two vowels. Thus, C2F24A has not lost the contrast between /y/ and /u/ completely. It could still be possible that she has merged in velar onset context while maintaining a coronal onset vs. labial onset contrast. Yet, there are other times when her token distributions overlap in the vowel space as already noted. It could be the more fronted articulations of /u/ and more retracted articulations of /y/ that other Cantonese speakers notice when they hear her speaking. C2F24A would, thus, be somewhere in the middle of the continuum between complete merger and complete distinction of /u/ vs. /y/. What is most certain is that C2F24A has brought the distribution of the two vowel classes into partial overlap. This is not seen in any of the GEN 1 or HK speakers.

To conclude, results presented in this section show no evidence of a merger of /y/ and /u/ on the group level. On the individual speaker level, however, there is clear evidence for variability in how distinct GEN 2 speakers produce the contrast between /y/ and /u/. For most speakers, the two vowels remain distinct, but for a few, the two have begun to merge in production and (possibly in perception as well). This merger is an innovation likely influenced by the local sociolinguistic context in Toronto since the merger is completely absent in GEN 1 and in HK data. In Section 6.5. I will provide further evidence showing that the facilitating factors are ones that point to contact with the English-speaking Toronto community.

6.4.2 Is there an increasing split between [i] and [ɪ] (Q2b)?

The second hypothesized change is an increasing split between [i] and [ɪ]. These two vowels are in complementary distribution in Cantonese, but Toronto English has phonetically similar counterparts (/i/ vs. /ɪ/) that are in contrastive distribution. Would Toronto English influence lead to an increasing acoustic distinction between these two Cantonese allophones? In a previous study (Tse 2016b), I showed evidence that this may be the case. I also discussed challenges in interpreting changes in this part of the vowel space without also incorporating acoustic data from Homeland speakers and from the English spoken by GEN 2 Toronto Cantonese speakers. For instance, [ɪ] retraction could be related to increasing phonetic distance between [i] and [ɪ] but [ɪ] retraction could also be an internally motivated change as well as a change influenced by the retracted /ɪ/ in Toronto English. With a larger set of data available including Homeland data, the results from the current study show that the increasing differentiation of [i] and [ɪ] is a change likely to have been initiated by GEN 1 speakers in Hong Kong. The results also show that Homeland speakers have advanced this split even further than has the GEN 2 group.

Pillai scores measuring the distance between [i] and [ɪ] were calculated for each speaker. Box plots showing the range of values for each of the three groups are shown in Figure 17 below. As illustrated in these box plots, the GEN 1 group has the greatest range of Pillai scores while the GEN 2 group has the smallest range. The GEN 1 group also has speakers with the lowest Pillai scores. This includes speakers with Pillai scores of less than 0.300. As a point of comparison, Hall-Lew (2009:150) defined 0.300 and below as merged in her study of the low-back vowel merger in San Francisco. Thus, in this range, the distinction between the two vowel classes is minimal. The highest Pillai scores within the GEN 1 group are between 0.600 and 0.700. Scores in this range typically suggest distinction. For example, in Hall-Lew's (2009) study of the low-back merger in San Francisco, the most distinct speaker had a Pillai Score of 0.709. If the change towards increasing distinction between [i] and [ɪ] reported in my earlier paper (Tse 2016b) were a change initiated by contact with Toronto English, we would expect GEN 2 speakers to have the highest Pillai scores. Yet, this is not the case. The wider range of scores within the GEN 1 group suggests that the increasing split in these two vowels was a change initiated in Hong Kong. Further supporting this is the range of Pillai scores from the Hong Kong group. Most speakers in the Hong Kong group have Pillai scores that are higher than for most GEN 1 speakers. In fact, the two lowest Pillai scores in the Hong Kong group come from the two oldest speakers. The difference between the HK group and each of the two Toronto groups is statistically significant according to an LSD Post-hoc test. Thus, it appears that the Hong Kong group is further advancing a change that was also brought over to Toronto by GEN 1 speakers.

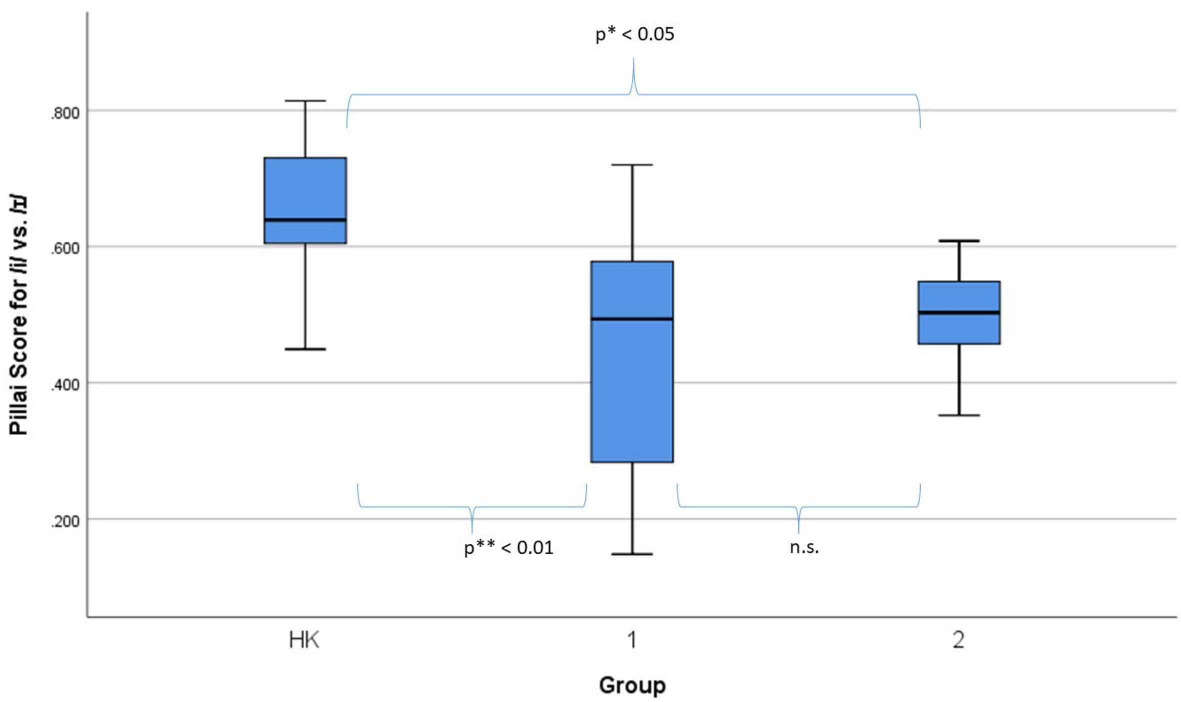

Figure 17. /i/ vs. /ɪ/ Pillai score range for each group

Variability in the acoustic distinction between [i] and [ɪ] is illustrated in the vowel plots that follow. Figure 18 shows the plot for C1M87A, the speaker with the lowest Pillai Score for the [i] vs. [ɪ] difference. For this speaker, the mean ± 1 SD ellipse for [ɪ] overlaps substantially with the ellipses for both [i] and [ɛ]. Illustrating the other end of the Pillai Score scale is the vowel plot for CXM20A, shown in Figure 19. This is the speaker with the second highest Pillai Score. In this plot, the acoustic distance between the ellipses for [i] and [ɪ] are far apart from each other. The ellipse for [ɪ], however, overlaps substantially with the ellipse for [ɛ]. All GEN 2 speakers are somewhere in between these two speakers in terms of acoustic distance between [i] and [ɪ]. The plot for C2M21C, the GEN 2 speaker with the lowest Pillai score, is shown in Figure 20. In this plot, the ellipses for [i] and [ɪ] show some overlap. For C2F24A, the GEN 2 speaker with the highest Pillai Score, [ɪ] overlaps with [ɛ] rather than with [i], as shown in Figure 21.

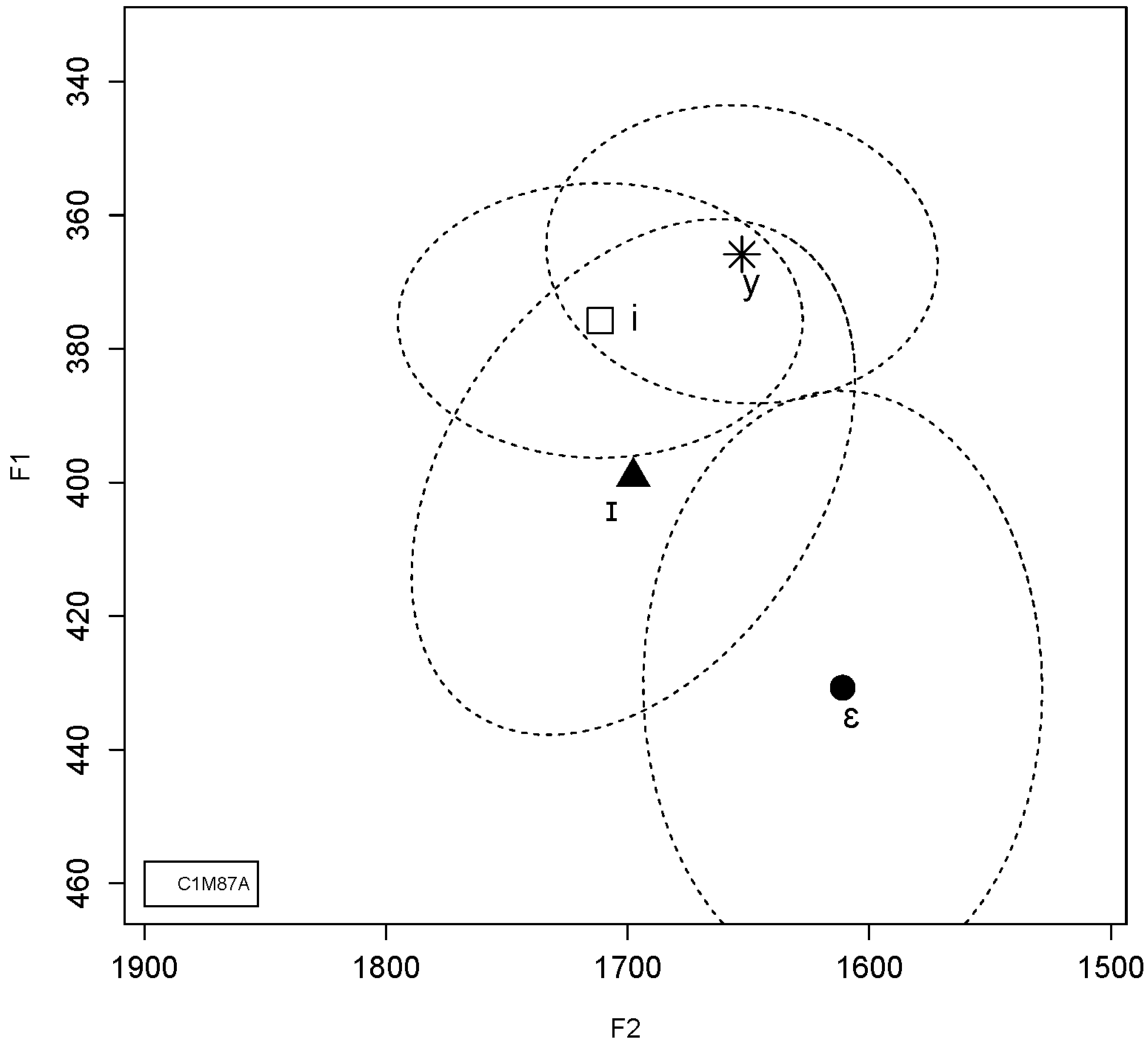

Figure 18. [i] vs. [ɪ] for C1M87A (Pillai Score: 0.148), Lobanov normalized formant values (Hz)

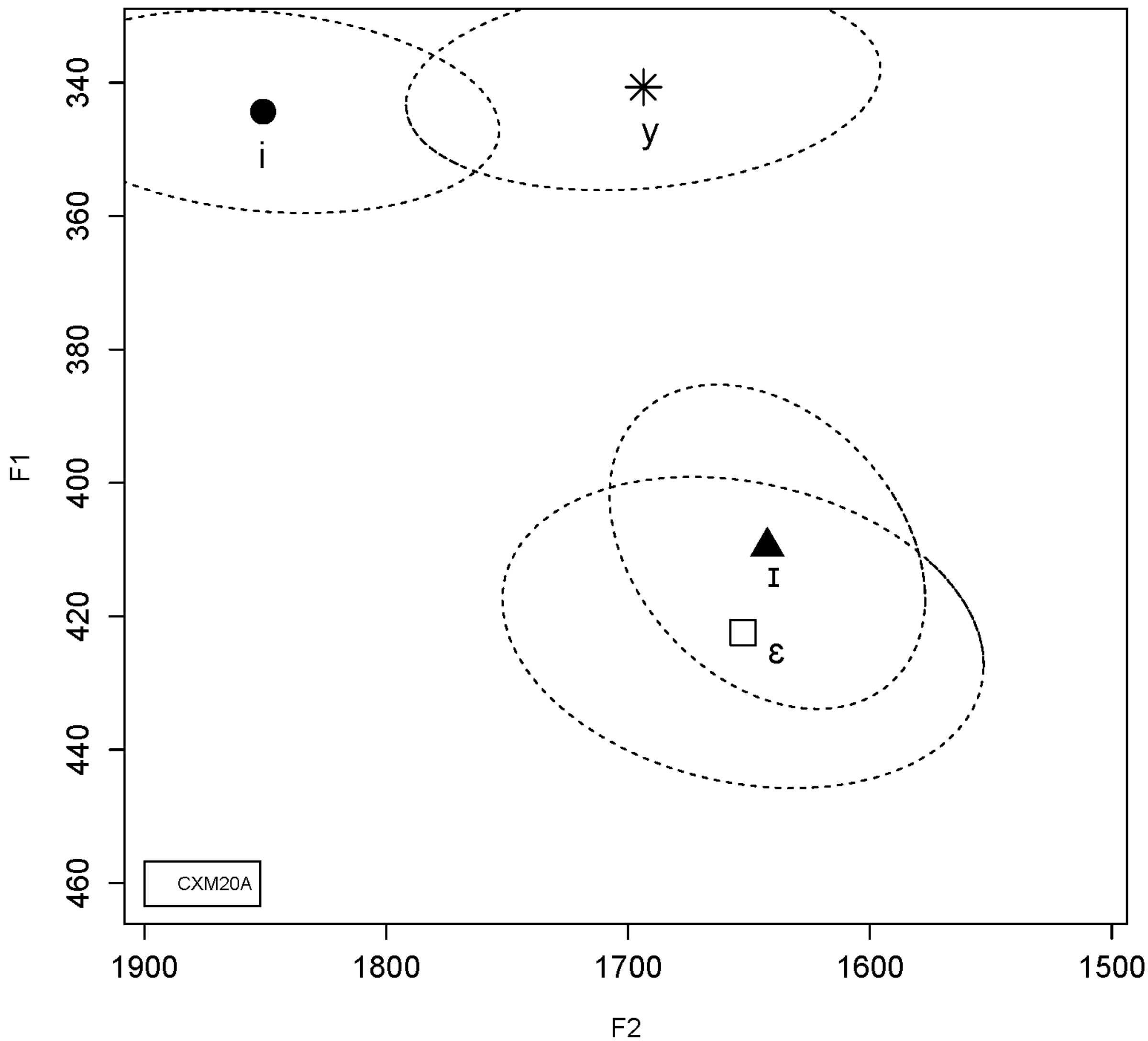

Figure 19. [i] vs. [ɪ] for CXM20A (Pillai Score 0.778), Lobanov normalized formant values (Hz)

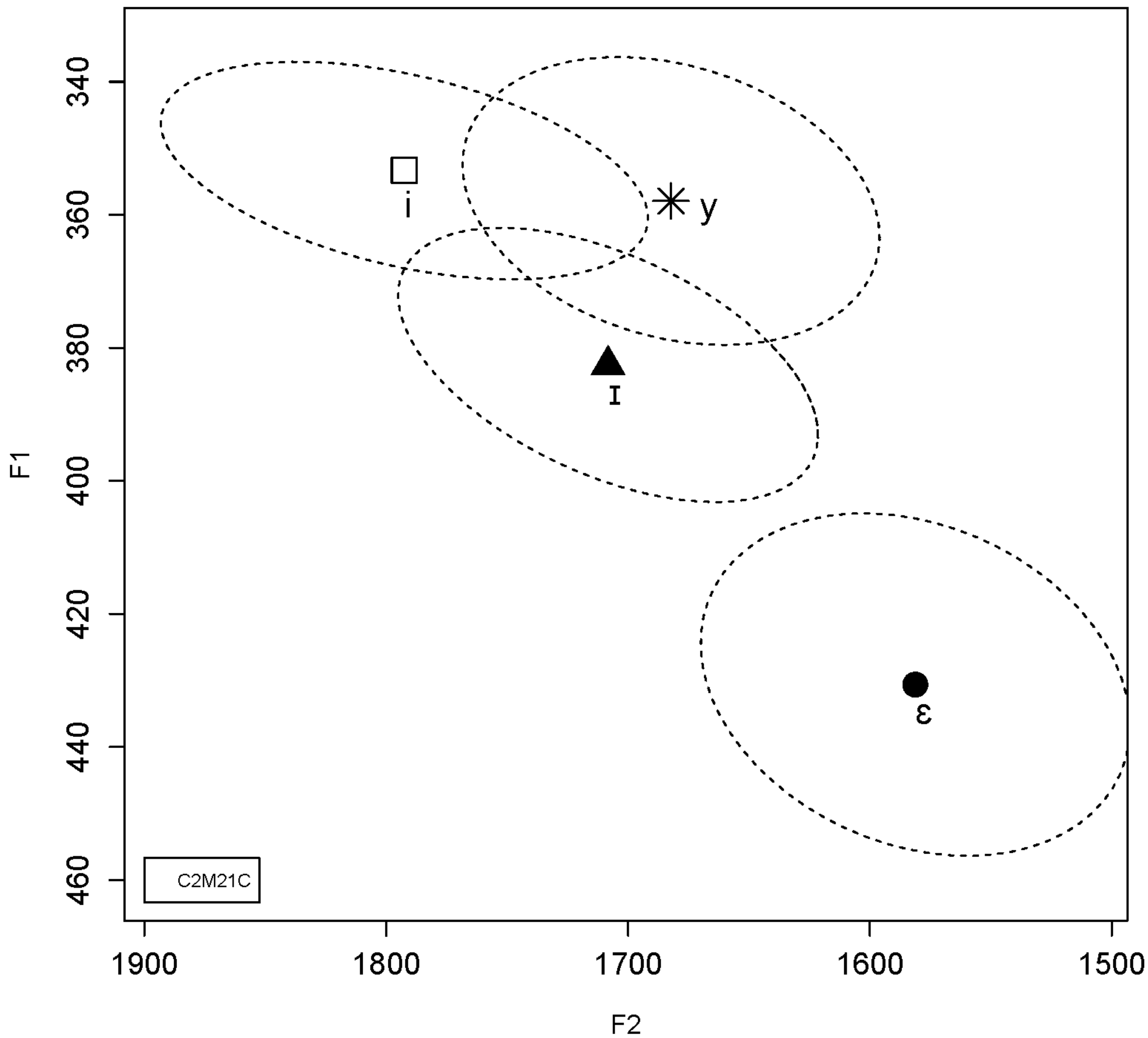

Figure 20. C2M21C (Pillai Score: 0.352) /i/ vs. /ɪ/ plot, Lobanov normalized formant values (Hz)

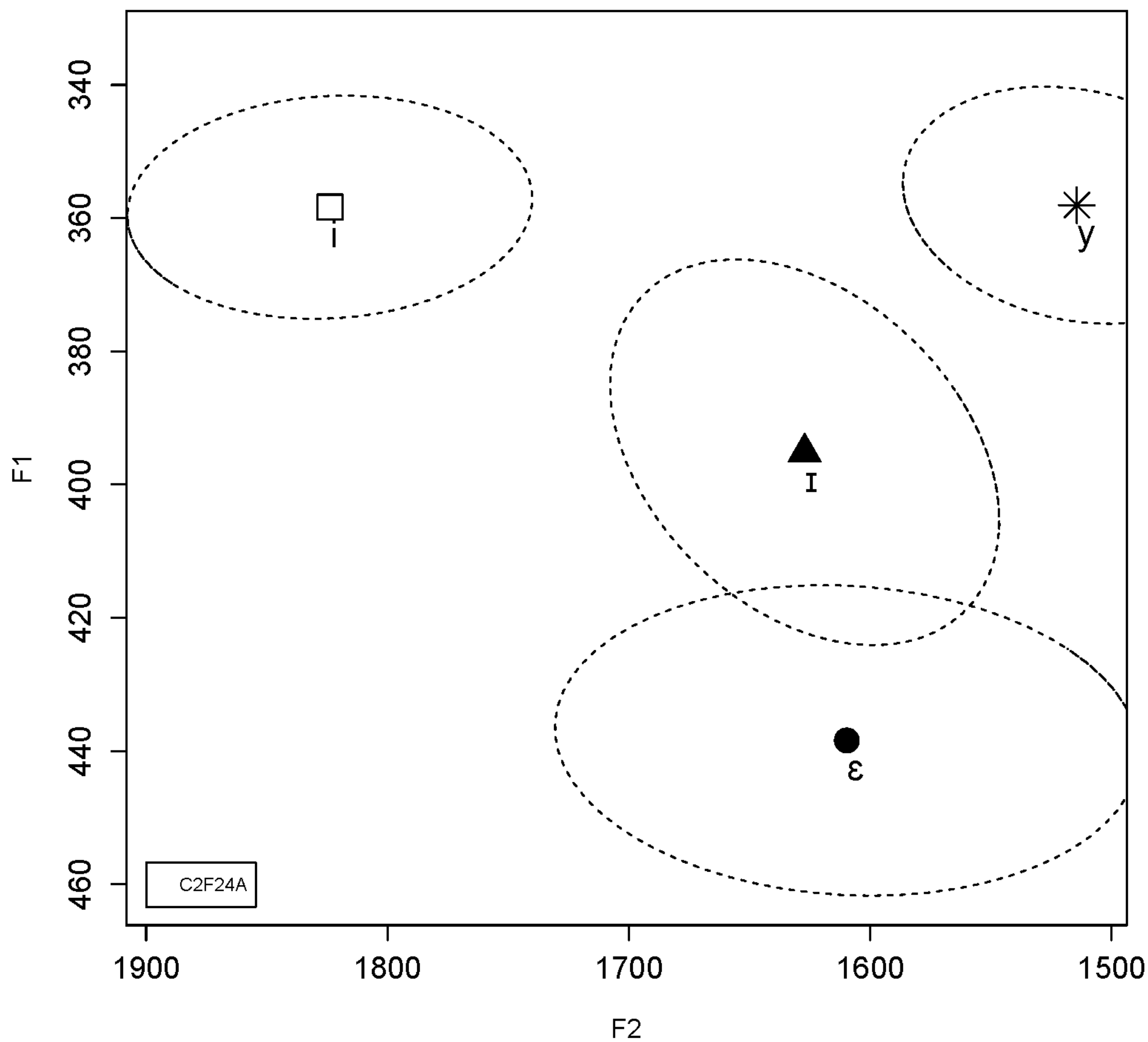

Figure 21. [i] vs. [ɪ] plot for C2F24A (Pillai Score: 0.608), Lobanov normalized formant values (Hz)

The results presented in this subsection reveal a puzzle. Why are GEN 2 speakers not further advancing a change started in Hong Kong? There are several possible explanations that involve the acoustic position of Cantonese [i] and [ɪ] in relation to their phonetically similar Toronto English counterparts. One possibility is that Cantonese [i] and [ɪ] for GEN 2 speakers have already merged with their production of Toronto English /i/ and /ɪ/. If GEN 2 speakers treat Cantonese [i] and [ɪ] as equivalent to Toronto English /i/ and /ɪ/, it would follow that they do not increase the distance between these two vowels. A second possibility is that GEN 2 speakers maintain a cross-linguistic difference by avoiding cross-linguistic merger of the Cantonese high front unround vowels with their phonetically similar Toronto English counterparts.

These possibilities would all require comparison with Toronto English data to address since it is uncertain where Cantonese [i] and [ɪ] are in relation to where Toronto English /i/ and /ɪ/ are in the acoustic space of GEN 2 speakers. Has Cantonese [ɪ] merged with the retracted /ɪ/ in Toronto English or do speakers avoid retraction to maintain a cross-linguistic distinction? This uncertainty underscores the importance of multiple comparisons in building accounts of contact-induced change (Nagy 2011). Further analysis of this part of the vowel space will have to wait until a future study that includes acoustic data from Toronto English. For the current study, what is clear is that GEN 2 speakers have not increased the F1/F2 distance between [i] and [ɪ].

6.4.3 Is there a split in /ɛ/ (Q2c)?

The third hypothesized change is an allophonic split in /ɛ/ based on velar context as was reported in my earlier work (Tse 2016a). With a larger set of data from the current study including additional contexts, the results presented below suggest that there may actually be two splits. This could be due to multiple sub-phonemic mappings between similar English allophones and Cantonese /ɛ/ based on phonetic context rather than a set of splits triggered by other changes in the vowel system as suggested in earlier work (Tse 2016a).

The results from the current set of data are consistent with previous results (Tse 2016a) in showing an inter-generational difference in coda consonant conditioning only along the F2 axis. Table 47 below shows results from three different models, one for each of the three groups. In each model, “speaker” and “word” were included as random effects while “coda context” was included as a fixed effect. F2 was the dependent variable in each model. “Coda” context is a significant predictor of F2 only for the GEN 2 group.

Table 47. Mixed effects model for the F2 of /ɛ/ for all three groups

<table>
<tr><td colspan="6">For all three models

Random Effects: “Speaker” and “Word”
Fixed Effect: “Coda Context”</td></tr>
<tr><td>GEN 1 data model</td><td colspan="4">GEN 2 data model
Coda Context (p = 0.00236)**</td><td>HK data model</td></tr>
<tr><td rowspan="5">No significant predictors</td><td></td><td>Coefficient (Hz)</td><td>Tokens</td><td>F2 Mean (Hz)</td><td rowspan="5">No significant Predictors</td></tr>
<tr><td>/ŋ/</td><td>39</td><td>258</td><td>1619</td></tr>
<tr><td>Open Syllable</td><td>-8</td><td>538</td><td>1564</td></tr>
<tr><td>/t/ or /k/</td><td>-30</td><td>40</td><td>1530</td></tr>
<tr><td colspan="4">r^2 [fixed] = 0.07, r^2 [random] = 0.388
r^2 [total] = 0.458</td></tr>
</table>

The results also suggest that the conditioning environment for F2 fronting would be better described as pre-nasal environment rather than pre-velar environment (as suggested in Tse 2016a). As shown in Table 47, the velar nasal, /ŋ/, conditions the highest F2 and hence the most fronted articulation of /ɛ/. The other velar consonant, however, is at the opposite end in conditioning the lowest F2 (and most retracted variants)[49]. This suggests that the conditioning environment involves manner of articulation rather than place of articulation. Thus, coda stops condition the lowering of F2 (retraction) while coda nasals condition higher F2 (fronting)[50]. Open syllable environment conditions neither fronting nor retraction. With fronting in one context, retraction in another, and neither retraction nor fronting in open syllable context, the results suggest the innovation of a nasal split and a coda stop split.

These splits could be influenced by allophonic similarity with Toronto English low and mid front vowels. The fronting of /ɛ/ in pre-nasal environment, for example, could be a pronunciation influenced by Toronto English /æ/ in pre-nasal context as in the word 'ban'. Toronto English, as in many dialects of North American English has a raised and fronted allophone of this vowel that is closer to IPA [ɛ] than to IPA [æ]. Cantonese /ɛ/ before stops may be more phonetically similar to Toronto English /ɛ/ in the same environment. This is because Toronto English /ɛ/ is lax and only occurs in closed syllable environment. This vowel is also a vowel involved with the Canadian Vowel Shift which results in the lowering and/or retraction of this vowel. The lower F2 in coda stop environment, thus, may be influenced by the retracted /ɛ/ in Toronto English.

The range of Pillai scores for the difference between open syllable and pre-nasal environment for each group is shown in the box plots in Figure 22. All but one GEN 1 speaker has a score of less than 0.300. As mentioned in Section 5.2.2, 0.300 and below is defined as

---

[49] Since there were only two tokens of coda /t/ and since these tokens patterned along with /k/, I ran a model in which /t/ and /k/ were included together as part of the same phonetic context. I also ran a separate model that included /t/ as a separate context from /k/. The results were virtually identical in terms of p-value and r^2 values.

[50] One factor I did not consider is interaction with other sound changes. The alveolarization of coda /ŋ/ and coda /k/ are well-documented sound changes in Hong Kong Cantonese (Matthews and Yip 2011:36–37). I impressionistically observed some GEN 2 speakers also participating in these changes. Even with the alveolarization of coda /ŋ/ and coda /k/, though, the conditioning environment for the split in /ɛ/ described here still holds.

merged in Hall-Lew (2009:150). Thus, most GEN 1 speakers do not show evidence of a pre-nasal /ɛ/ split. In contrast, the GEN 2 group has the widest range of scores from a low of 0.057 to a high of 0.535. Almost half of the GEN 2 speakers have a Pillai score above 0.300, which suggests evidence of an allophonic split. LSD Post-hoc tests show that the difference between GEN 1 and GEN 2 is significant ($p < 0.05$). Examples of vowel plots for speakers with the lowest Pillai scores are shown in Figure 23 (C1F78A, Pillai Score: 0.039) and in Figure 24 (C2F22A, Pillai Score: 0.081). In these plots and in the rest of the plots in this section, tokens in open syllable context are represented as empty squares, while tokens in nasal context are represented as dots. Triangles represent tokens of coda /k/. The ellipses indicate one standard deviation from the mean. Both Figure 23 and Figure 24 show substantial overlap across all three contexts and lack of phonetically conditioned variation.

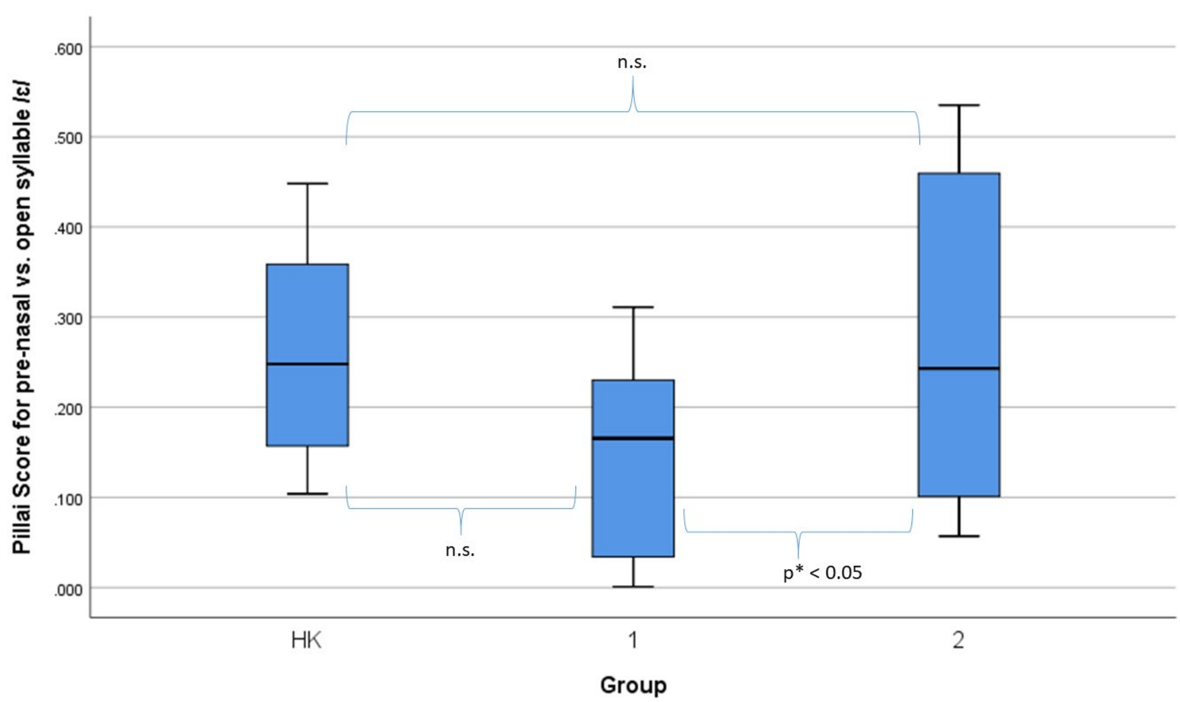

Figure 22. Boxplots showing range of Pillai Scores for open syllable vs. pre-nasal /ɛ/

Although the difference in Pillai Score ranges between the GEN 1 and GEN 2 groups is significant, Figure 22 also shows that the difference between the GEN 2 and the Homeland groups is not significant. At first glance, this appears to suggest that the Homeland group may be undergoing the same change as the GEN 2 group. Upon closer inspection, however, the non-significance of the Pillai Score range differences is due to inclusion of F1 in the calculation of Pillai Scores. In Table 48, Table 49, and Table 50 are models of F1 variation of /ɛ/ for each group. They all show that velar coda context has a significant raising effect on F1. This means that for all three speaker groups, /ɛ/ is articulatorily lower in the vowel space when preceding a velar consonant. Thus, the fact that the Pillai Score variation range between the GEN 2 and Homeland groups is not significantly different could be due to the fact that both groups have a split along the F1 axis. Nevertheless, it remains clear based on the analysis above that the two groups can still be described as different from each other along the F2 axis. Thus, if there is contact-induced change in this vowel, the influence of Toronto English appears to be limited to F2. In the vowel plots for GEN 2 speakers that follow, the split is primarily along the F2 axis.

Table 48. Mixed effects model for the F1 of /ɛ/ for GEN 1

Random Effects ($r^2 = 0.239$) Speaker, Word			
Fixed Effect ($r^2 = 0.114$) Coda Context ($p = 1.51e-07$)***	Coefficient (Hz)	Tokens	F1 Mean (Hz)
Velar (/k/ and /ŋ/)	10	431	436
Open Syllable	-10	703	420
r^2 [total] = 0.353			

Table 49. Mixed effects model for the F1 of /ɛ/ for GEN 2

Random Effects ($r^2 = 0.261$) Speaker, Word			
Fixed Effect ($r^2 = 0.060$) Coda Context ($p = 0.000421$)***	Coefficient (Hz)	Tokens	F1 Mean (Hz)
Velar (/k/ and /ŋ/)	9	296	439
Open Syllable	-9	538	429
r^2 [total] = 0.321			

Table 50. Mixed effects model for the F1 of /ɛ/ for the Homeland group

Random Effects ($r^2 = 0.250$) Speaker, Word			
Fixed Effect ($r^2 = 0.246$) Coda Context ($p = 3.08e-09$)***	Coefficient (Hz)	Tokens	F1 Mean (Hz)
Velar (/k/ and /ŋ/)	15	200	431
Open Syllable	-15	346	417
r^2 [total] = 0.496			

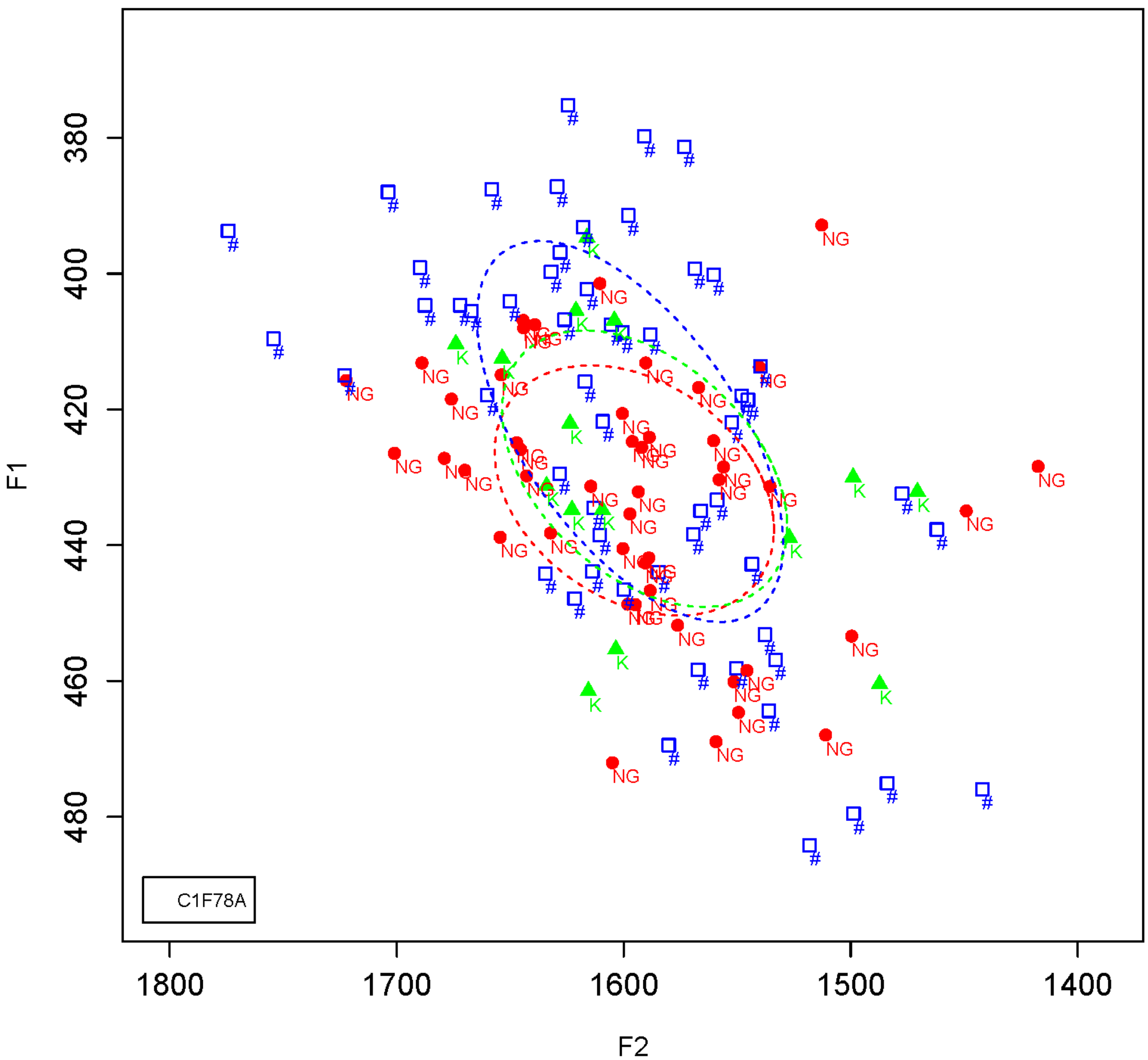

Figure 23. Plot for C1F78A (Pillai Score: 0.039) with /ɛ/ tokens in three contexts (open syllable in squares, nasal coda in dots, stop codas in triangles), Lobanov normalized values (Hz)

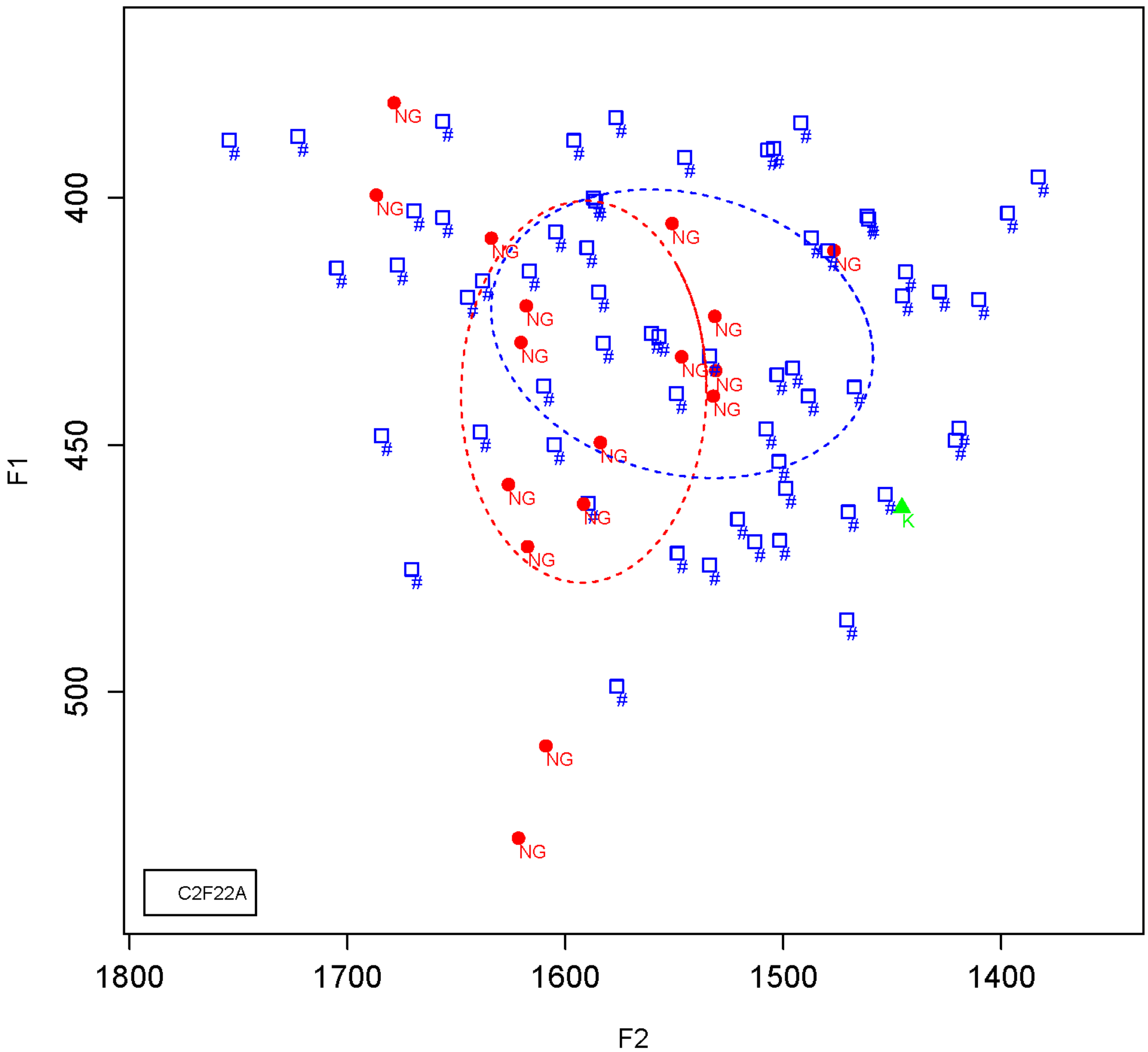

Figure 24. Plot for C2F22A (Pillai Score: 0.081) with /ɛ/ tokens in three contexts (open syllable in squares, nasal coda in dots, stop coda in triangle), Lobanov normalized values (Hz)

Figure 25 shows a vowel plot for C2F24A. This is the speaker with the highest Pillai score. For this speaker, there is a clear distinction based on coda context. The standard deviation ellipses do not overlap at all. The next vowel plot, shown in Figure 26, is for C2M22A, the speaker with the second highest Pillai score. Nasal coda context is also clearly separated from open syllable context, but unlike for C2F24A, the nasal coda context tokens are lowered in addition to being fronted.

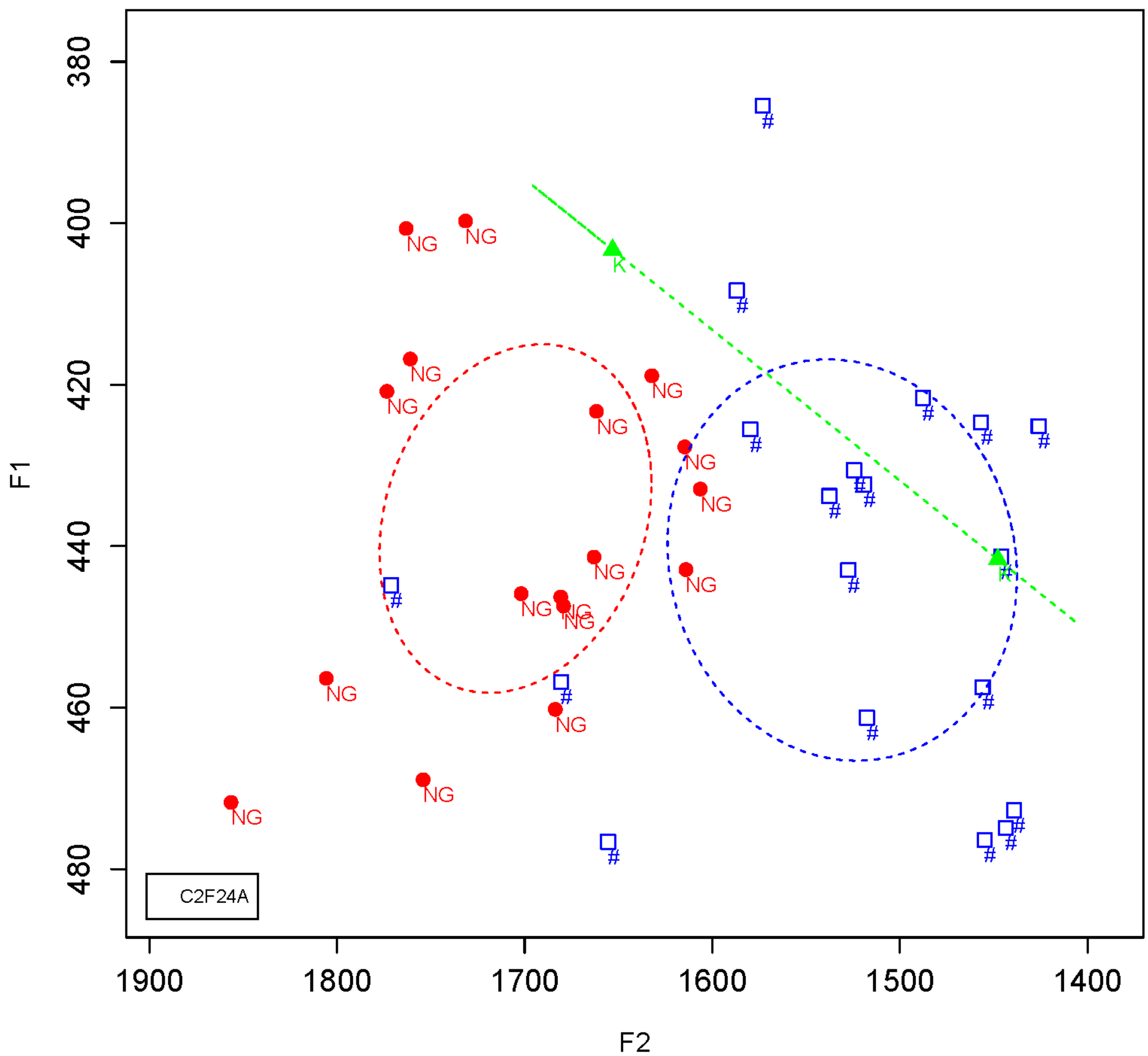

Figure 25. Plot for C2F24A (Pillai Score: 0.535, highest in data) with /ɛ/ in 3 contexts (open syllable in squares, nasal coda in dots, stop codas in triangles), Lobanov normalized values (Hz)

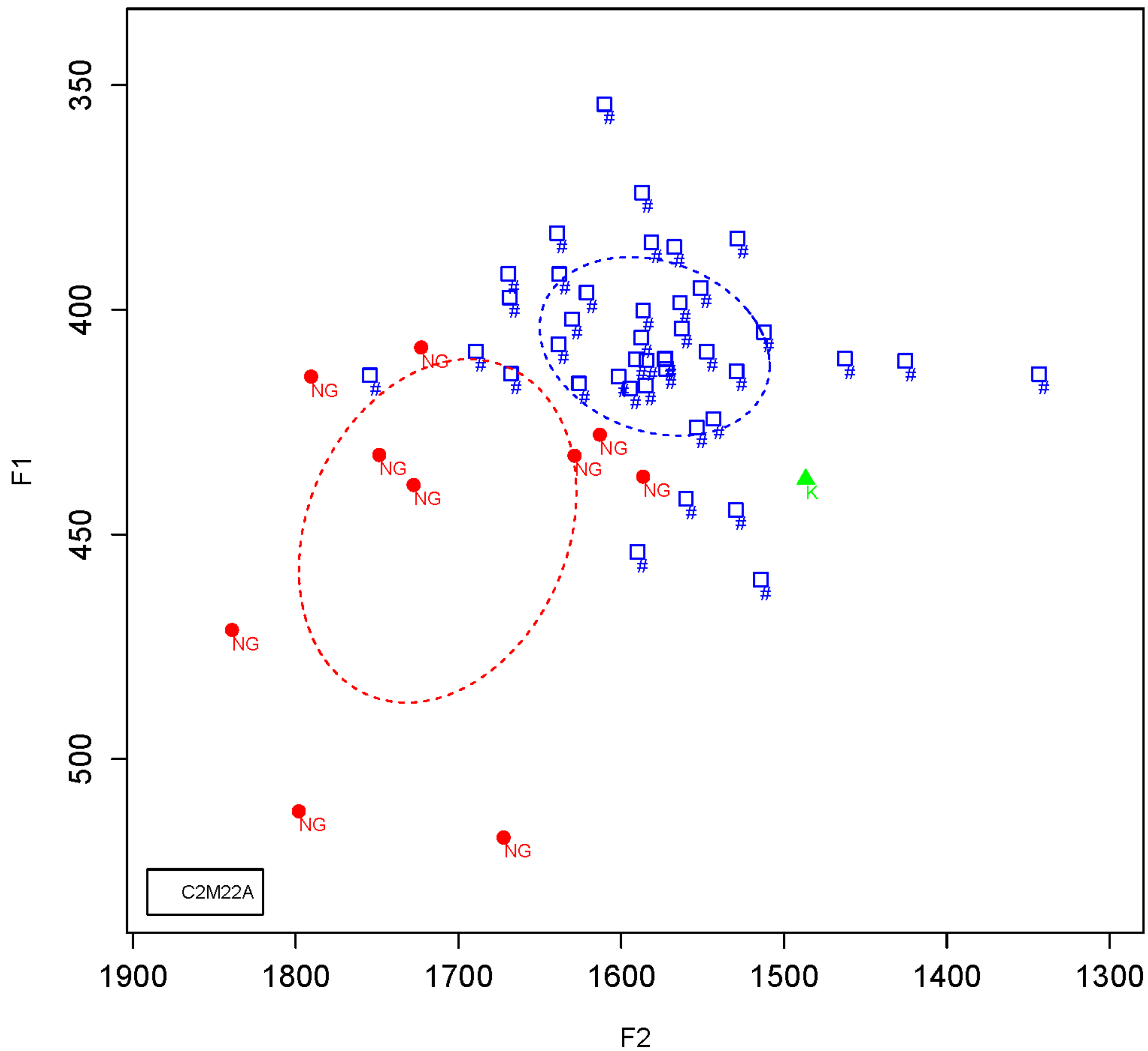

Figure 26. Plot for C2M22A (Pillai Score: 0.517) with /ε/ in three contexts (open syllable in squares, nasal coda in dots, stop coda in triangle), Lobanov normalized values (Hz)

Evidence for the innovation of a split based on stop coda context is not as strong as the evidence for the innovation of a split based on nasal coda context. This is because coda stop environment is not well represented in the data analyzed. Many speakers had only a few tokens in this context while two speakers had zero tokens. Coda context tokens, however, were not universally retracted among GEN 2 speakers. For example, there are two tokens of coda /k/ in the plot for C2F24A (Figure 25). Only one of these two tokens is retracted, but only slightly. Both C2F22A (Figure 24) and C2M22A (Figure 26), on the other hand, have one token each of coda /k/. Both tokens are retracted, but it is difficult to say whether or not this retraction pattern holds with a larger number of tokens for the same speakers. The only speaker that clearly shows the two different splits is C2F41A (Figure 27).

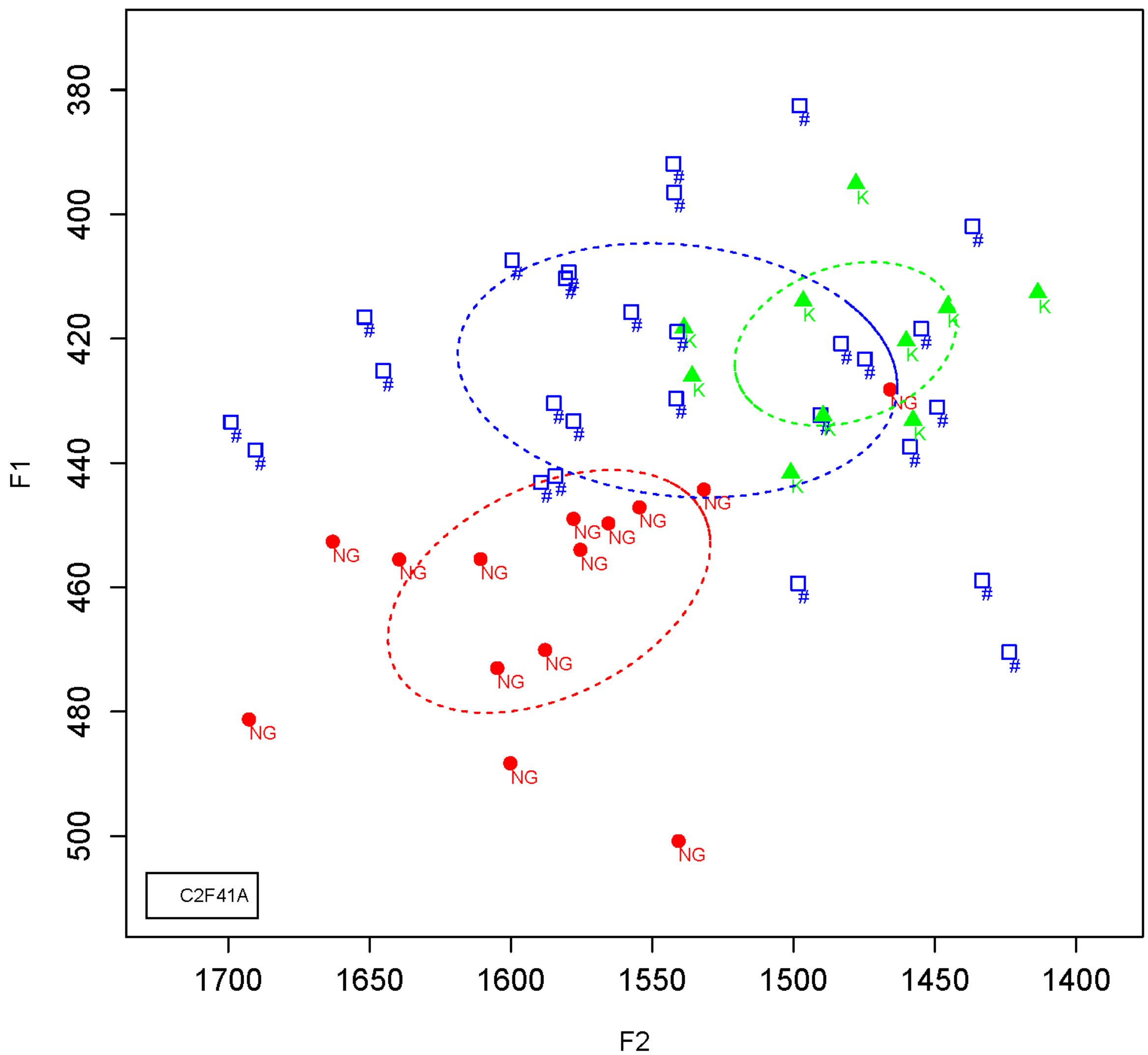

Figure 27. Plot for C2F41A with /ɛ/ in three contexts (open syllable in squares, nasal coda in dots, stop codas in triangles), Lobanov normalized values (Hz)

To conclude this section, I have shown evidence for the GEN 2 innovation of a split based on nasal coda context and a split based on stop coda context. Pre-nasal variants favor fronted variants while stop coda environment favors retraction. Open syllable environment favors F2 values between these two extremes. I have also proposed a link between these variants and phonetically similar allophones in Toronto English. Coda context was not a statistically significant predictor of F2 variation for either the GEN 1 group or the Homeland group. These two splits, thus, appear to be Toronto innovations. In Section 6.5, I will show additional evidence that the nasal split is influenced by contact with Toronto English. The coda stop split will not be investigated further due to overall lack of tokens and lack of tokens for individual speakers.

6.4.4 Split in /ɔ/ (Q2d)?

Finally, the last hypothesized change I discuss is whether there is a split in /ɔ/ conditioned by velar context as reported in an earlier study, which involved a small sample size of 15 tokens of /ɔ/ per speaker (Tse 2016a). With a larger sample of data that includes more phonetic environments and an average of more than 100 tokens per speaker, results from the current study show that the /ɔ/ split may be better described as one conditioned by an open vs. closed syllable distinction. Results also show that the split is not a GEN 2 innovation as previously reported. This highlights the importance of considering all possible phonetic contexts when describing vowel variation.

The F1 of /ɔ/ as the dependent variable was modeled separately for GEN 1 and for GEN 2. Both models included "speaker" and "word" as random effects and "coda consonant" as the fixed effect. Results from the GEN 1 group are shown in Table 51. Coda context came out significant ($p < 0.001$)[51]. The model included five possible values for "coda" context. The coefficient values and the means are very similar for /k/, /n/, /ŋ/, and /t/. These environments can

---

[51] I also ran separate models with "syllable type" (open vs. closed) instead of specific coda segment. These models had almost the same r^2 values but higher AIC values. I present the models with specific coda segment to show the pattern in detail.

all be described as closed syllable context. Open syllable context conditions the lowest F1. Thus, there appears to be a split based on open vs. closed syllable context. The model for the GEN 2 group is shown in

Table 52. The coefficient values and means are also very similar to each other for /k/, /n/, /ŋ/, and /t/. Open syllable context also conditions the lowest F1 as it does for GEN 1.

Table 51. Mixed effects model for the F1 of /ɔ/ for GEN 1

Random Effects ($r^2 = 0.233$) Speaker, Word			
Fixed Effect ($r^2 = 0.148$) Coda Consonant ($p = 2.66e-18$)***	Coefficient (Hz)	Tokens	F1 Mean (Hz)
/k/	6	360	451
/n/	5	37	451
/ŋ/	4	660	450
/t/	3	18	446
# (Open syllable)	-18	1220	429
r^2 [total] = 0.381			

Table 52. Mixed effects model for the F1 of /ɔ/ for GEN 2

Random Effects ($r^2 = 0.246$) Speaker, Word			
Fixed Effect ($r^2 = 0.136$) Coda Consonant ($p = 2.33e-09$)***	Coefficient (Hz)	Tokens[52]	F1 Mean (Hz)
/t/	6	3	460
/k/	6	317	449
/ŋ/	5	333	449
/n/	0	20	450
#	-17	1050	426
r^2 [total] = 0.382			

[52] Three tokens for coda /t/ is not a typo. This is also the least common context for GEN 1 speakers as shown in Table 51.

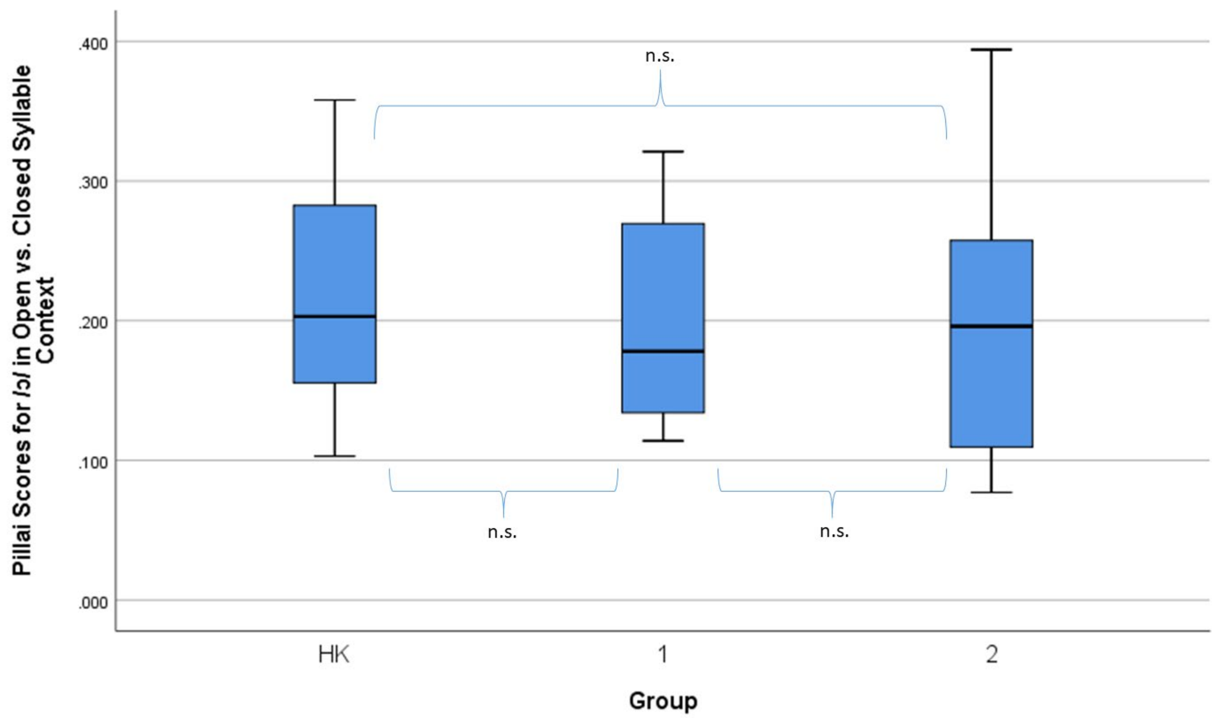

Figure 28. Range of Pillai Scores (/ɔ/ in open vs. closed syllable context) across three groups

Since both GEN 1 and GEN 2 show a split in /ɔ/, this split cannot be described as a GEN 2 innovation. Furthermore, the models shown in Table 51 and in

Table 52 appear to be very similar. This suggests lack of inter-generational change. Further evidence of the lack of inter-generational change in this part of the vowel system can be found in the similarity in the range of Pillai scores for /ɔ/ tokens in open vs. closed syllable context. The range of Pillai scores for each group is shown in Figure 28 below. Most speakers across all three groups have Pillai scores of less than 0.300. LSD Post-hoc tests showed no significant inter-group differences. Thus, there is lack of evidence for change in this part of the vowel system.

6.5 EXTERNAL FACTORS? (Q3)

The third research question addresses the extent to which demographic, ethnic orientation, or language amount factors (Sex, EOQ Score, EOQ responses, Cantonese % Score, Cantonese WC Score, English WC Score) can account for the innovation of shifts, mergers, and splits observed. The GEN 2 innovations discussed below include the retraction of /y/, the incipient merger of /y/ and /u/, and the nasal split in /ɛ/ along the F2 axis. The increasing split in [i] vs. [ɪ] and the split in /ɔ/ were found not to be GEN 2 innovations and are thus not further discussed in this section.

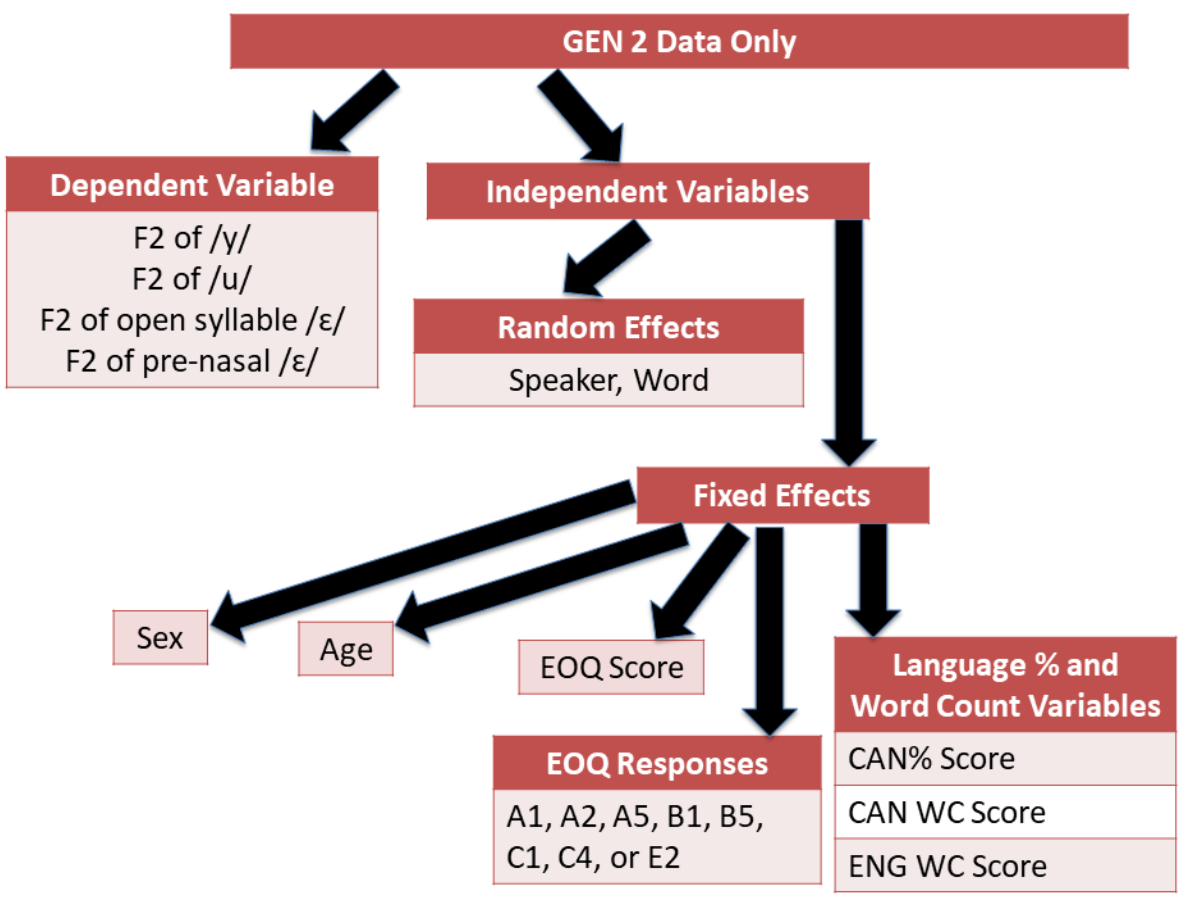

Figure 29. Factors considered in mixed effects modeling of GEN 2 innovations

The first sub-section below addresses both /y/ retraction and the incipient merger of /y/ and /u/. It includes results showing how these two changes are inter-related to each other. The second sub-section focuses on the nasal split in /ɛ/. Figure 29 is a diagram showing all the language external factors considered in the mixed effects models included in this section. All models include "speaker" and "word" as random effects. Up to five fixed effects were considered in each model. They include "age", "sex", "EOQ Score", an EOQ response variable, and a language use variable. These variables were explained in detail in Chapter 5. There were eight possible EOQ response variables and three possible language use variables. No more than one of each was included in the same model due to co-linearity.

6.5.1 Accounting for the retraction of /y/ and the merger of /y/ and /u/

The results presented below suggest that the retraction of /y/ (shown in mixed effects modeling of the F2 of /y/) and the incipient merger of /y/ and /u/ (shown by examining vowel plots for speakers with the lowest Pillai Scores) are related changes. Although no inter-generational difference in the F2 of /u/ was identified in Section 6.3, two of the factors accounting for variation in the F2 of /u/ within the GEN 2 group were factors that also came out significant in modeling intra-group variation in the F2 of /y/. These two factors were the ENG WC Score and the CAN % Score. Factors that did not come out significant in any model for the F2 of /y/ or for the F2 of /u/ include Sex, Age, EOQ Score, and EOQ questions B1, B2, B3, B4, B5, C2, C4, and E2.

Based on r^2 values, the models that include ENG WC Score both accounted for a larger percentage of variation than the models that include CAN % Score. The model for the F2 of /y/

is shown in Table 53 while the model for the F2 of /u/ is shown in Table 54. For the F2 of /y/, there is an inverse relationship between the ENG WC Score and the F2 of /y/. A higher ENG WC Score correlates with lower F2 (more retracted variants of /y/). For the F2 of /u/, the relationship between ENG WC Score and F2 is a positive correlation. Thus, higher ENG WC Scores correlate with higher F2 of /u/ (more fronted articulations). Putting the two together, higher ENG WC Scores predict both /y/ retraction and /u/ fronting. The overall effect of these two processes is a trend towards merger of /y/ and /u/. This shows that those GEN 2 speakers who used more English WC in their interviews lead in merging the two vowels. The inter-generational retraction of /y/ (shown in Section 6.3) and the incipient merger of /y/ and /u/ among a few GEN 2 speakers (shown in Section 6.4.1) are, thus, related processes.

Table 53. GEN 2 mixed effects model for the F2 of /y/

Random Effects (r^2 = 0.315) Speaker, Word			
Fixed Effect (r^2 = 0.100) ENG WC Score (p = 7.17e-03)*		Coefficient	Tokens
continuous	+1	-7.141 Hz	351
r^2 [total] = 0.415, AIC = 4079			

Table 54. Mixed effects model for the F2 of /u/

Random Effects (r^2 = 0.199) Speaker, Word			
Fixed Effect (r^2 = 0.182) ENG WC Score (p = 6.09e-03)*		Coefficient	Tokens
continuous	+1	8.479 Hz	165
r^2 [total] = 0.381, AIC: 1962			

The two models that include CAN % Score also came out significant. Although these two models have lower r^2 values, they also have lower AIC. They also paint the same general picture as do the models that include ENG WC Score. Table 55 shows a positive correlation between CAN % Score and the F2 of /y/. This means that those who used the most Cantonese (measured in terms of token percentages) tended to have higher F2 of /y/ (more fronted articulations). Those who used less Cantonese would, thus, produce the most retracted variants of /y/. Since those who used less Cantonese also tended to have higher ENG WC Scores, this would point to the same relationship between language use and F2 retraction of /y/ shown in the model presented in Table 53. Similarly, in

Table 56, the model of the F2 of /u/ shows that those who used more Cantonese in the interviews tended to have more retracted variants of /u/. Fronting of /u/ would thus be led by those who used less Cantonese in their interviews according to this model. This model, thus, shows a similar relationship between language use and F2 of /u/ as shown in the model presented in Table 54. The overall effect is the same. Merger of /u/ and /y/ is led by those with lower CAN % Scores.

Table 55. GEN 2 mixed effects model for the F2 of /y/ with CAN % Score

Random Effects (r^2 =0.3174) Speaker, Word			
Fixed Effect (r^2 = 0.0946) CAN % Score (p = 6.91e-03)*		Coefficient	Tokens
continuous	+1	187	351
r^2 [total] = 0.412 AIC 4073			

Table 56. GEN 2 mixed effects model for the F2 of /u/ with CAN % Score

Random Effects ($r^2 = 0.234$) Speaker, Word			
Fixed Effect ($r^2 = 0.123$) CAN % Score ($p = 0.0163$)*		Coefficient	Tokens
continuous	+1	-204 Hz	165
r^2 [total] = 0.357 AIC 1957			

Table 57. GEN 2 mixed effects model of the F2 of /y/ with EOQ C1

Random Effects ($r^2 = 0.3429$) Speaker, Word			
Fixed Effect ($r^2 = 0.0791$) Vowel ($p = 0.0398$)*	Coefficient	Tokens	F2 Mean (Hz)
Cantonese (2)	45	155	1625
Both (1)	17	171	1601
English (0)	-63	25	1529
r^2 [total] = 0.422			

Table 58. GEN 2 mixed effects model of the F2 of /u/ with EOQ C1

Random Effects ($r^2 = 0.172$) Speaker, Word			
Fixed Effect ($r^2 = 0.163$) C1 ($p = 4.13e-03$)***	Coefficient	Tokens	F2 Mean (Hz)
English (0)	86	19	1288
Cantonese (2)	-34	51	1163
Both (1)	-52	95	1147
r^2 [total] = 0.335			

Aside from the language use scores, EOQ Question C1 ("What language does your family speak when you get together?") also came out significant for both the F2 of /y/ and the F2 of /u/, but only in models in which C1 is the only fixed effect included. The model with C1 for the F2 of /y/ is shown in Table 57 while the model with C1 for the F2 of /u/ is shown in Table 58. Those who said that they use English when they speak with their family are the ones who favor the most retracted variants (lower F2) of /y/ and the most fronted (higher F2) variants of /u/. Once again, these models show the same relationship between language use and innovation as do the models with CAN % Score and ENG WC Score. As reported in Table 26 in Section 5.3.2, only two GEN 2 speakers answered "English" for this question. These two speakers are also the speakers with the two lowest CAN % Score and the two highest ENG WC Scores.

EOQ A1 and EOQ A5 also came out significant for the F2 of /y/ but not for the F2 of /u/. EOQ A1 and EOQ A5 both relate to ethnic identity. The model for EOQ Question A1 ("Do you think of yourself as Chinese, Canadian, or Chinese-Canadian?) is shown in Table 59. This model shows that the lowest F2 values (most retracted) of /y/ occur with those who answered "Canadian". Those who responded "Chinese" or "Chinese-Canadian" had the most fronted tokens of F2. This also supports a contact-based explanation to the retraction of /y/ since we would expect those who identify more strongly with Canadian culture to show more English influence in their Cantonese speech.

The results for the model with EOQ A5 ("when you were growing up, were the kids in your school Chinese?") are shown in

Table 60. According to this model, those who said that most of their peers in school were Chinese are the ones who produce the most retracted variants of /y/ (lower F2). At first glance, this seems to suggest the opposite of what the model of EOQ A1 shows. If those who self-identify as Canadian produce the most retracted variants, we would expect those who grew up

with more non-Chinese peers to also produce the most retracted variants. The model of EOQ A5, however, shows the opposite relationship.

One possible explanation for these results is that GEN 2 speakers who were exposed to a more diverse group of peers were able to better learn how to differentiate between English and Cantonese phonology. As shown by the unanimous responses to EOQ B3 ("do you prefer to speak Cantonese or English?") discussed in Chapter 5, most GEN 2 speakers prefer English regardless of the ethnic makeup of the schools in which they attended. Those who went to school with more Chinese peers may have had less of a need to learn how to differentiate between English and Cantonese phonology. Attending a school with mostly ethnic Chinese peers in the Toronto area does not necessarily mean that all Cantonese speaking students prefer to talk to each other in Cantonese. They may maintain Cantonese as a language used primarily at home but prefer English in a school context because English is the dominant language of instruction. Having more Chinese peers while being in schools in which English is the primary language of instruction could mean becoming more English-dominant over time and having English phonology affect Cantonese phonology. With fewer non-Chinese peers, there may be less of a need to distinguish between the phonologies of these two languages.

Table 59. GEN 2 mixed effects model of the F2 of /y/ with EOQ A1

Random Effects ($r^2 = 0.216$) Speaker, Word			
Fixed Effect ($r^2 = 0.188$) EOQ A1 ($p = 1.84e-04$)***	Coefficient	Tokens	F2 Mean (Hz)
"Chinese" (2)	37	111	1629
"Chinese-Canadian" (1)	26	148	1631
"Canadian" (0)	-63	92	1541
r^2 [total] = 0.404			

Table 60. GEN 2 mixed effects model of the F2 of /u/ with EOQ A5

Random Effects ($r^2 = 0.3964$) Speaker, Word			
Fixed Effect ($r^2 = 0.0951$) EOQ A5 ($p = 0.0256$)*	Coefficient	Tokens	F2 Mean (Hz)
"Both Chinese and non-Chinese" (1)	37	148	1631
"Non-Chinese" (0)	9	136	1609
"Chinese" (2)	-46	67	1550
r^2 [total] = 0.422			

6.5.2 Accounting for the nasal split in /ɛ/

As was the case for /y/ retraction and /u/ fronting, both ENG WC Score and CAN % Score are shown to be significant predictors in accounting for the nasal split in /ɛ/. Unlike for the F2 of /y/ and for the F2 of /u/, however, none of the EOQ question variables came out significant. Sex, Age, and EOQ Score also did not come out significant.

The nasal split is modeled based on variation in the F2 of /ɛ/ in pre-nasal context. This is possible because of the stability of /ɛ/ in open syllable context. This was shown in models of F1 and F2 that include only GEN 2 tokens of /ɛ/ in open syllable context. Not a single significant predictor of F1 or F2 variation in /ɛ/ in open syllable context was identified. Models of F1 variation of /ɛ/ in pre-nasal context also showed lack of significant predictors. All of this suggests stability of /ɛ/ in open syllable context and movement of /ɛ/ primarily on the F2 axis. The nasal split can, thus, be modeled based only on the F2 of /ɛ/ in pre-nasal context.

Table 61 shows a model with ENG WC Score as a fixed continuous effect. This model shows a positive correlation between the F2 of pre-nasal /ɛ/ and ENG WC Scores. This means that those who lead in the fronting (and in the nasal split) of /ɛ/ are those who used the most English vocabulary in the interview samples. The model with CAN % Score shows a similar relationship. This model, shown in Table 62, shows an inverse relationship between CAN % score and the F2 of pre-nasal /ɛ/. Those with lower CAN % Scores, thus, produce more fronted variants of pre-nasal /ɛ/. Those who lead in pre-nasal /ɛ/ fronting are, thus, the ones who use less Cantonese and more English in the interviews. As was the case for the F2 of /y/ and for the F2 of /u/, all significant models point to contact with Toronto English as the source of innovation. To summarize this sub-section, results for the pre-nasal split in /ɛ/ show the same language use factors involved in the incipient merger of /y/ and /u/. The next chapter will provide a more detailed discussion of how these language use factors play a role in accounting for these changes.

Table 61. Mixed effects model for the F2 of [ɛŋ]

Random Effects (r^2 = 0.339) Speaker, Word			
Fixed Effect (r^2 = 0.163) ENG Vocabulary Score (p = 1.89e-03)***		Coefficient (in Hertz)	Tokens
continuous	+1	-161	258
r^2 [total] = 0.502 AIC 2911			

Table 62. Mixed effects model for the F2 of [ɛŋ]

Random Effects (r^2 = 0.373) Speaker, Word			Tokens
Fixed Effect (r^2 = 0.122) CAN % Score (p= 7.22e-03)***		Coefficient (in Hertz)	
continuous	+1	-161	258
r^2 [total] = 0.495 AIC 2906			

6.6 RESULTS SUMMARY

Based on the results presented in this chapter, we can address the three research questions. To reiterate, the questions are Q1) Is there evidence for inter-generational vowel shifting? 2) Is there evidence for vowel mergers or vowel splits influenced by contact with Toronto English? 3) To what extent can demographic, ethnic orientation, or language use factors (Sex, EOQ Score, specific EOQ responses, Cantonese % score, Cantonese vocabulary score, English vocabulary score) account for the specific shifts, mergers, and splits observed in the data?

The answer to the first question is that there is evidence for only one vowel showing inter-generational shift. The results presented in this chapter show that /y/ is significantly more retracted (lower F2) among the GEN 2 group than it is among the GEN 1 group. For the rest of the vowel system, there is overall maintenance. In contrast, the Homeland group shows evidence for four different vowel shifts in apparent time. One of these changes, the fronting of /i/, is also an apparent time change in Toronto. The retraction of /y/, however, is not a change in Homeland Cantonese.

For the second question, results from this chapter show evidence that some speakers have an incipient merger of the two high round vowels. Results also show evidence for the innovation of two allophonic splits for the vowel /ɛ/. This vowel is fronted preceding nasals and retracted preceding stops. For the contrast between /i/ and /ɪ/, results show a change in apparent time initiated by GEN 1 speakers that involves increasing the acoustic distance between these two vowels. This change has advanced even further in the Homeland group. The GEN 2 group, however, does not seem to be advancing this change any further. Results also show lack of evidence of change in /ɔ/. Although GEN 1 speakers do show a split based on open vs. closed syllable context, GEN 2 speakers also have the same split. The GEN 2 group does not appear to be changing this vowel in any way.

Finally, the third question focuses on what factors could account for shifts, mergers, and splits identified in the GEN 2 data. Results show that ENG WC Score and CAN % Score account for the retraction of /y/, the incipient merger of /y/ and /u/, and the pre-nasal split in /ɛ/. These factors both point to contact with Toronto English and with Toronto or Canadian culture as the source of these innovations. EOQ C1, which is about home language, also came out significant for both the retraction of /y/ and the fronting of /u/, but for the pre-nasal /ɛ/ split. This further supports language use factors as the factors that best account for innovations within the GEN 2 group. Two EOQ questions related to ethnic identity also came out significant for /y/ retraction, but not for changes in other vowels. The results for EOQ A1 suggest that those with a stronger sense of Canadian identity are the ones that lead in /y/ retraction. The other EOQ factor that came out significant was about ethnic composition of the schools in which GEN 2 speakers attended. Results for this factor suggest a more complicated relationship between language use and identity. What is clear from the overall results is that language use factors can consistently account for all of the GEN 2 innovations identified. Ethnic orientation may be related to language use, but language use is also the product of many other variables rather than a deterministic outcome of one's ethnic orientation and peer network.

The next chapter will provide a more detailed discussion of how to interpret all of these results. They appear to complicate Labov's (2007) Transmission and Diffusion model.

7.0 DISCUSSION

The results presented in the previous chapter show that GEN 2 Toronto speakers are not a homogeneous group. Early bilingualism in a community provides access to a greater variety of linguistic resources than available in monolingual communities. This makes possible seemingly oppositional outcomes such as the lack of change in the production of Cantonese vowels for some speakers and for others, the innovation of structural changes influenced by Toronto English. In Section 7.1, I discuss the role of social meaning in preserving cross-linguistic phonetic differences and in preventing contact-induced vowel shifts. In Section 7.2, I discuss how the few structural changes that have been observed are related to CAN % and ENG WC Scores as well as relative token frequency of particular phonetic contexts in spontaneous speech. I discuss how these findings challenge the TD Model (discussed in Section 2.4) in Section 7.3. I will then discuss the theoretical implications of these findings for models of contact-induced change in Section 7.4 (as discussed in Section 3.1). Although both the TK and VC Models avoid problematic assumptions inherent in the TD Model, I will argue in Section 7.5 that an even better model would be one that combines the most important insights of both the TK and VC Models. Finally, I conclude in Section 7.6 by discussing the limitations of the current study and future research directions.

7.1 THE SOCIAL MEANING OF AVOIDING VOWEL SHIFTS

Given experimental studies showing heritage speakers producing two phonetically distinct systems, as discussed in Section 3.3, it should not be surprising to find lack of contact-induced vowel shifts. Furthermore, as discussed in Section 4.3, Cantonese has become the language of many local businesses such as those present in the "Chinatowns" found within the city limits and in several suburbs. What this means is that there is social support for language maintenance Cantonese (and other heritage languages) in the Greater Toronto Area. Thus, for both psycholinguistic and macro-level social reasons, it should not be surprising to find lack of changes in the overall vowel system for GEN 2 speakers. In this section, I argue that micro-level social reasons also play a role in accounting for the lack of vowel shifts in Toronto Heritage Cantonese. I will support this argument my presenting excerpts from HLVC Project interviews in which speakers discuss cross-linguistic differences they see between Cantonese and English. What these excerpts suggest is that GEN 2 speakers find it socially meaningful to maintain two distinct phonological systems even though many of these speakers acknowledge that their English speaking skills are superior to their Cantonese speaking skills.

In the first interview excerpt shown in Table 63 below, we can see a specific example of how knowledge of GEN 1 pronunciation patterns helps GEN 2 speakers avoid merging Cantonese vowels with phonetically similar Toronto English counterparts. In this excerpt, the interviewer asks C2M21C if he could tell where someone is from based on the way they speak. C2M21C hesitates for a while and then comes up with an example involving the pronunciation of "Loblaw's", the name of a local supermarket chain.

Table 63. Excerpt from C2M21C

Original (Cantonese and English)	English Idiomatic Translation
Oh, hou2 jung4 ji6 ze3, "Loblaws" [lɑblɑs]. Right? ze1 hai6 go2 go3, go3 gaan1 aa3 grocery store. jyu4 gwo2 nei5 hai6 jing1 man2 hai6 nei5 wui5 ho2 nang4 gong2 zau6 hai6 "Lob-", uh, "Low-blaws" [lowblɑs], or "Lob-laws" [lɑblɑs], right? Like, go2 di1 zung1 man2 le1, nei5, jyu4 gwo2 nei5 ze1 hai6 jau5 hou2 daai6 go3 daai6 go3 hau2 jam1 le1, nei5 zau6 wui5 teng1 dou2, pei3 jyu4, uh "Lob-laws" [lowplɔs] aa3, zau6 si4 uh, "Lub-law" [lɐpla]. (laughs) nei5 zau6 zi1 dou3 hai6, zik6 cing5 hai6 hai2 bin1 dou6 lei4, keoi5 aa3.	Oh, easy, "Loblaws" [lɑblɑs], right? Like that, that grocery store. If you are English speaking, you would probably say [lowb …] [lowblɑs] or [lɑblɑs] right? Like, for Chinese speakers, if like you've got someone with a big, a big accent, you would hear, for example, uh [lowplɔs] or maybe uh [lɐpla] … [laughing] … then you can tell exactly where someone is from.

In Figure 30 is a vowel plot for C2M21C that shows the F1 and F2 of both the first and second syllable of his pronunciation of "Loblaw's". This plot also includes vowel ellipses to show the typical range of his pronunciation of Cantonese /ɔ/, /ɐ/, and /a/. Both syllables of his first token (intended to reflect the Toronto English pronunciation) are outside the range of these three vowels. They are both pronounced exactly where we would expect Toronto English /ɑ/ to be pronounced. Thus, there is no question about C2M21C's ability to produce the Toronto English vowel as we would expect other Torontonians to pronounce it. His second and third tokens of "Loblaw's" were intended to be two different pronunciations that might be used among native English speakers. The first is [lowblɔs], which appears to be the British pronunciation while the second is [lɑblɑs], the Toronto English pronunciation. The British pronunciation appears to be exactly where we would expect the /ow/ and /ɔ/ of UK English to be. In the second Toronto English pronunciation, the [ɑ] in the first syllable is below the range of Cantonese /ɔ/ as expected, but the second token is within the range of Cantonese /ɔ/, though still relatively low in the articulatory vowel space. The third and fourth tokens in the excerpt were intended to be two different ways that a GEN 1 or Homeland speaker would pronounce the name. The first pronunciation is [low.plɔs], while the second pronunciation is [lɐp.la]. Both vowels in the first Cantonese pronunciations are within the vowel ellipses shown in Figure 30. The [ɐ] in the second Cantonese pronunciation is outside the ellipse for /ɐ/, but Cantonese /ɐ/ has been described as retracted preceding labial segments (Bauer and Benedict 1997:71)[53]. This pronunciation is, thus, within the range of what we would expect for the GEN 1 pronunciation. The two Cantonese variants also show two different strategies for dealing with /p/ + /l/ sequences, which are otherwise absent in native vocabulary. As Bauer and Benedict (1997:379) mention, some Cantonese speakers have no problem with adopting clusters with /l/ in English loan words. Thus, [low.plɔs] is one possible Cantonese pronunciation while the other, [lɐp.la], has the [p] and [l] split across two syllables. Cantonese syllables with coda stops can only have a monophthong in the nucleus. Thus, the vowel is [ɐ] rather than [ow] in the form [lɐp.lɑ]. The second vowel, however, impressionistically seemed to be more like a hybrid pronunciation between the Cantonese vowel /ɔ/ and Toronto English /ɑ/. This was also immediately followed by laughter. It could be that the Cantonese pronunciation he attempted to imitate is not his normal (or "vernacular" following Labov's term) pronunciation, thus, he laughed knowing it was not what he is used to. At the same time, he clearly recognizes the different pronunciations of "Loblaw's" based on one's language background.

As discussed in Section 6.4.4, there is no evidence for inter-generational change for /ɔ/. In contrast, the Hong Kong group is lowering this vowel in apparent time. This is perhaps coincidentally the same direction of movement expected if Cantonese /ɔ/ were being influenced

[53] This is based on earlier impressionistic observations. Although not reported in the current study, this observation appears to be an accurate description of GEN 1 vowel phonology based on acoustic data as well.

by Toronto English /ɑ/. The fact that /ɔ/ is not lowering among GEN 2 speakers even though it could as it has in Hong Kong suggests that GEN 2 speakers as a group could be avoiding lowering to maintain a cross-linguistic distinction between Cantonese /ɔ/ and Toronto English /ɑ/.

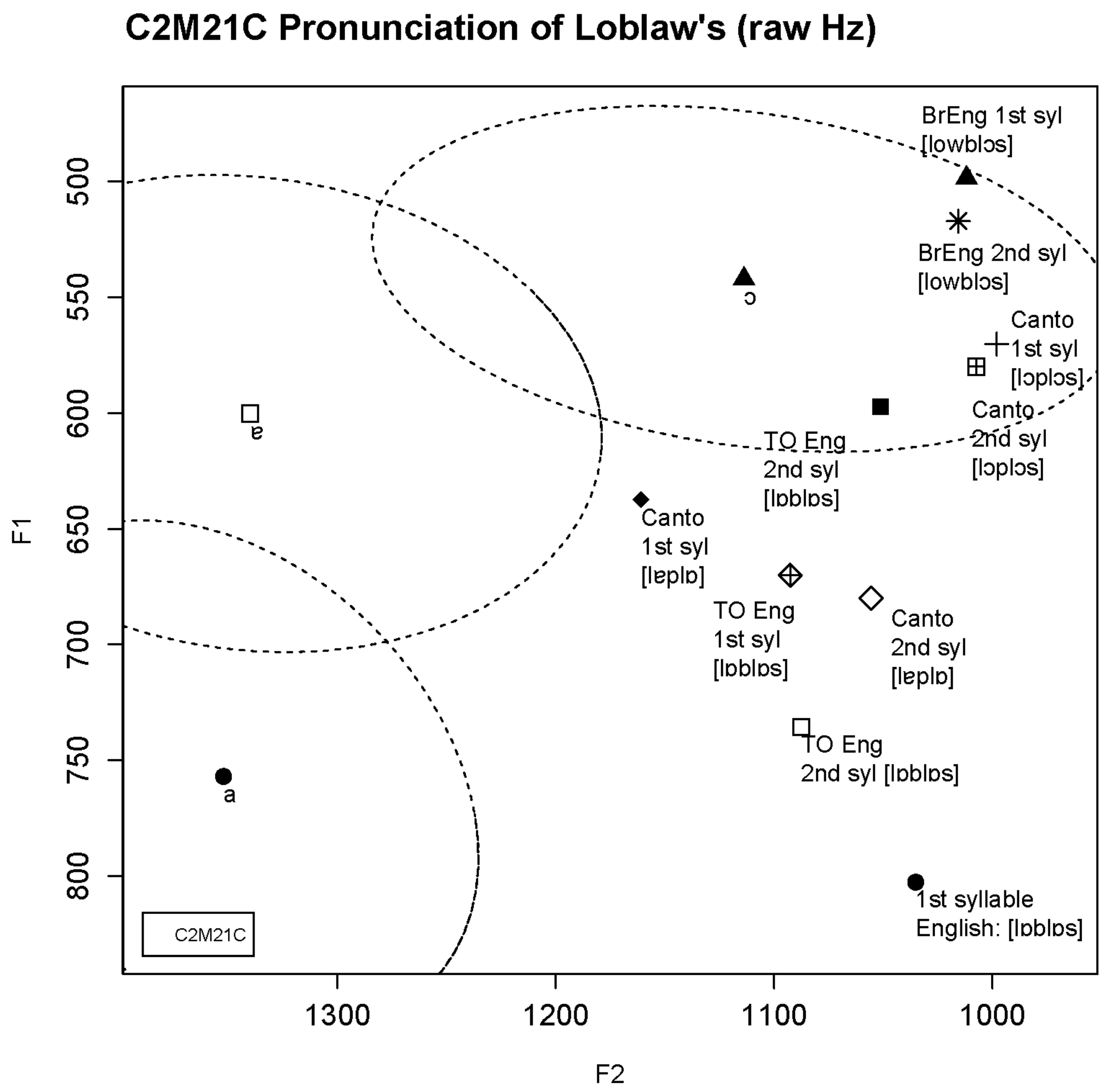

Figure 30. C2M21C Pronunciation of <Loblaw's>

The acoustic evidence presented above makes it clear that at least one GEN 2 speaker is able to produce cross-linguistic distinctions between phonetically similar vowels. C2M21C also shows evidence of knowledge of Cantonese syllable structure, when he pronounces <Loblaw's> as [lɐp.la]. Furthermore, the metalinguistic commentary about pronouncing the name of a supermarket chain shows that the distinct pronunciations are socially meaningful. For C2M21C, there are two different ways in which native English speakers pronounce "Loblaw's" and two different ways in which Cantonese speakers who speak English with a "big accent" would pronounce this name. If GEN 2 speakers in general are able to produce and recognize a distinction between the /ɔ/ in Cantonese and the /ɑ/ in the Toronto English pronunciation of "Loblaw's", then it should not be a surprise that GEN 2 speakers have not lowered Cantonese /ɔ/. The same reason likely explains the lack of vowel shifts for other vowels.

Table 64. Interview excerpt from C2M27A

	Original (Cantonese and English)	**English Idiomatic Translation**
Interviewer	Gam2 jyu4 gwo2 jau5 jan4 tung4 lei5 gong2 "You sound like you're from Hong Kong" um, "by the way you talk" do you think it's an insult or a compliment?	So, if someone says to you, "You sound like you're from Hong Kong um by the way you talk", do you think it's an insult or a compliment?
C2M27A	I think it's a compliment.	I think it's a compliment.
Interviewer	Dim2 gaai2 le1?	Why?
C2M27A	Jan1 mai6 o5 gok3 dak1 o5 zung1 man2 dou1 m4 hai6 gei2 hou2 gam2 zau6 jyu4 go2 waa6, jan4 dei6 waa6 hoeng1 gong2 hai6 ... sound like Hoeng1 Gong2 di1 zung1 man2 zau6 it's, they're really good. ze1 hai6 di1 keoi5 dei6 go2 di1 zung1 man2 hai6 hou2 hou2 aa3.	Because I think that my Chinese is not even that great so if someone says … someone says Hong Kong … sounds like Hong Kong Chinese, then, it's … they're really good, like their Chinese is really really good.
Interviewer	Gam2 jyu4 go2 lei5 gong2 gan2 jing1 man2 le1, gam2 keoi5 dei6 waa6 lei5 hai6 Hoeng1 Gong2 lei4?	Now, if you're speaking English, and then they say you must be from Hong Kong?
C2M27A	Insult	Insult

Evidence also suggests that GEN 2 speakers place high value on speaking different languages without a perceptible accent in any of them. The social value of "accentless" speech in both English and Cantonese could also explain the lack of vowel shifts. This sentiment is illustrated in an excerpt from the interview for C2M27A shown in Table 64. C2M27A thinks it would be a compliment if someone says that his Cantonese sounds like the Cantonese from someone from Hong Kong. At the same time, he would find it insulting if someone says that his English sounds like he is from Hong Kong. For GEN 2 Cantonese speakers, Chinese-Canadian identity does not mean speaking Cantonese with an English-influenced accent nor does it mean speaking English with a Cantonese-influenced accent. It means being able to speak either language in appropriate contexts without an accent influenced by the other language.

Awareness of two distinct phonological systems as well as social value in speaking Cantonese without an accent, however, are only part of the picture. Excerpts from several interviews show evidence that the knowledge that GEN 2 speakers have of two distinct phonological systems is also productive. Many GEN 2 speakers appear to use correspondence rules as a strategy for dealing with lexical gaps in their knowledge of Cantonese. Thomason (2007) describes the use of correspondence rules as based on knowledge (either conscious or unconscious) of the sound correspondences that exist in these different languages. When these speakers overextend their knowledge to new vocabulary or even to existing vocabulary in one of these languages based on their knowledge of existing sound correspondences, these speakers are using correspondence rules. Thomason says that the use of correspondence rules provides "excellent evidence of bilingual speakers' ability to manipulate equivalent forms, usually phonological, in their two languages" (2007:46).

Many examples of correspondence rules come from speakers of genetically related languages. Nurse and Hinnebusch (1993:269), for example, describe the use of correspondence rules among speakers of Bajuni (also known as Tikuu) who also speak Standard Swahili. Both languages belonging to the Bantu family. These speakers recognize a sound correspondence in which /c/ in Bajuni corresponds with Standard Swahili /t/ across a large set of cognates. They have extended their knowledge of this sound correspondence to new words that they have learned from Standard Swahili. For example, the word for "tape" in Standard Swahili is [tepu], while the word for "team" is [timu]. Both of these are English loanwords that have entered Standard Swahili. When Bajuni speakers speak Bajuni, however, they pronounce these words as [cepu] and [cimu] respectively. This is not about the inability of Bajuni speakers to pronounce /t/ but about Bajuni speakers making productive use of their knowledge of how two languages are related to each other. Bajuni speakers have, thus, introduced the words [cepu] and [cimu] by using a correspondence rule that equates /t/ in Standard Swahili vocabulary to /c/ in Bajuni. They would still say [tepu] and [timu], however, when speaking Standard Swahili.

The Toronto Cantonese case shares some similarities to the Bajuni case discussed by Nurse and Hinnebusch (1993:269). Both cases involve English loan words entering a second language. In the Toronto Cantonese case, GEN 2 speakers are familiar with how GEN 1 speakers nativize English loan words. In Table 65 is an excerpt from C2F22A that illustrates the metalinguistic awareness that is typical of GEN 2 speakers. Here C2F22A is describing how she sees the speech of Hong Kong Cantonese speakers.

Table 65. Excerpt from C2F22A

Original (Cantonese and English)	English Idiomatic Translation
It's like riddled with English, right? Like keoi5 dei6 zi6 gei2 zing2 go2 di1 jing1 man4 zi6 so like ci1 si2, uh, like "cheese" and like kuk1 kei4 beng2, which is cookies, go2 di1 je5 m4 hai6 zung1 man4 zi6 lei4 ge3 so2 ji5 … Yeah.	It's like riddled with English, right? Like they themselves create those English words so like "ci1 si2", uh like "cheese" and like "kuk1 kei4 beng2", which is cookies, those things are not even Chinese words, so … yeah

In this excerpt, C2F22A identifies two specific English loan words that have entered the Hong Kong (and GEN 1) speaker lexicon. She has no trouble pronouncing these words with the correct tones and segmental features and in the way that GEN 1 speakers would pronounce them. She also has no trouble pronouncing the English source words using Toronto English phonology. Since C2F22A learned both Cantonese and English at an early age, there are neither cognitive nor articulatory constraints on her ability to pronounce either the Cantonese words (with Cantonese phonology) or the English source words (with Toronto English phonology). Like other GEN 2 speakers, she sees the use of English loan words pronounced with Cantonese phonology as a characteristic of the speech of Hong Kongers.

As I show in some of the excerpts below, the pervasiveness of English loan word vocabulary in the Cantonese of GEN 1 speakers has given GEN 2 speakers enough exposure to be able to generalize patterns in how to convert an English word into a Cantonese word. Without explicit instruction, they have mastered the rules of loanword phonology as discussed in studies such as Silverman (1992). While Thomason (2007) has discussed how the use of correspondence rules can lead to deliberate change in the lexicon of a language, in the Cantonese case, it leads to deliberate non-change in Cantonese phonology even if this means use of more English vocabulary while speaking Cantonese.

In the excerpt below taken from the picture naming task, C2F24A is unsure of the Cantonese word for “balloon”. After hesitating for a moment, she applies correspondence rules to come up with a word. This involves converting the iambic stress pattern (unstressed + stressed) in the English word “balloon” to a Tone 3 (mid-level) + Tone 1 (high-level) pattern (see Silverman 1992 for discussion of stress to tone mapping in Cantonese loan word phonology). The unstressed vowel in “balloon” corresponds with Cantonese /a/ while the vowel in the second syllable corresponds with Cantonese /u/.

> C2F24A: baa3 lun1 … [laughs]. It’s true. Cantonese is just like add a fobby accent and like [tæk1 si2], <dik1 si2>, right?

Here we can see that correspondence rules are used as a strategy for dealing with a gap in the speaker’s lexical knowledge of Cantonese. In reflecting on her spontaneous decision to say <baa3 lun1>, C2F24A says somewhat jokingly that Cantonese involves adding a “fobby accent” to English words. The word “fob” comes from the abbreviation “F. O. B.”, meaning “fresh off the boat”. This is a common expression used within many Asian North American communities to refer to recent immigrants. C2F24A, thus, sees the use of English loan words pronounced with Cantonese phonology as part of what it means to speak Cantonese. The most essential part of speaking Cantonese for her is the use of Cantonese phonology. Her comment reflects her awareness that this is common behavior for GEN 1 speakers when they speak Cantonese.

GEN 2 speakers also recognize that correspondences between English source words and loan words that have entered Cantonese have exceptions (usually due to influence from the writing system or borrowing via other varieties of Chinese as explained in Section 4.2.3). For instance, the first time C2F24A says “taxi” in the excerpt above, she follows the stress to tone correspondence rules (Tone 1 + Tone 2 corresponding to English trochaic stress, see Silverman 1992, Bauer and Benedict 1997:395), but leaves the English segments intact including the [æ], which is not a vowel present in Cantonese. She then self-corrects by saying the actual Cantonese word, which is <dik1 si2>.

In the excerpt from C2M21B, shown in Table 66, we see evidence of correspondence rules in suprasegmental phonology as a strategy for dealing with lexical and structural gaps in knowledge of Cantonese. The transcription includes Tone numbers for English words uttered using Cantonese suprasegmental phonology. These tones follow English stress to Cantonese tone correspondences as described by Silverman (1992).

After C2M21B says "more1 re3-laxed1 laa1" (with Tone 3 + Tone 1 corresponding to iambic stress in English disyllabic words, see Silverman 1992:302), he switches to English across all linguistic levels (phonology, lexicon, and syntax). He then comments on how he felt like he was speaking Cantonese even though much of what he said was English vocabulary and syntax pronounced with Cantonese phonology. He then repeats the last phrase he says with the Cantonese tonal patterns: "more1 re3-laxed1 laa1". C2M21B turns out to be the speaker that has the second lowest CAN % Scores among all of the GEN 2 speakers interviewed. Here we see that this appears to be due to weak vocabulary in Cantonese. In spite of his limited Cantonese vocabulary, this excerpt demonstrates that C2M21B's strongest area in Cantonese is productive use of Cantonese phonology. What this example clearly shows is that GEN 2 speaker knowledge of GEN 1 phonology is likely stronger than GEN 2 speaker knowledge of GEN 1 vocabulary and morpho-syntax. With knowledge and productive use of GEN 1 phonology, it should, thus, not be a surprise to see lack of evidence for contact-induced vowel shifts.

Table 66. Excerpt from C2M21B

	Original (Cantonese and English)	English Idiomatic Translation
Interviewer	Gam2 aa3 lei5 baa4 baa1 maa1 mi4 jim6 m4 jim6? Strict.	So, your parents, are they strict?
C2M21B	Strict ah.. maybe as a kid, ah siu2 siu2, strict. Ah... but as ngo5 grow1 up1 hm... more1 re3laxed1 laa1. [clears throat] I like how I say my English words with the Chinese accent and I make you think, I, I think that's it's Chinese.	Strict uh, maybe as a kid, a little bit strict. Uh … but as I grew up hm … they became more relaxed. [clears throat] I like how I say my English words with the Chinese accent and I make you think, I, I think that it's Chinese.
Interviewer	More1 re3laxed1.	More relaxed (in Cantonese phonology).
C2M21B	More1 re3laxed1 laa1.	More relaxed (in Cantonese phonology).
Interviewer	Yeah, that's all that counts	Yeah, that's all that counts.
C2M21B	It's all zung1 man2 there.	It's all Chinese there.

The use of correspondence rules is so widespread and productive that it can even mislead GEN 2 speakers into thinking a particular word is part of the repertoire of loan words used by GEN 1 and Homeland speakers. To illustrate, C2F21B tells a story about asking her step brother what the Cantonese word for "stubborn" is. He told her it was "si6 daa1 baan3". This follows the stress to tone correspondences in which primary stress in English corresponds to Tone 1 (high level) in Cantonese and unstressed syllables correspond to Tone 3 (mid-level). It also follows the pattern of splitting consonant clusters with an epenthetic syllable pronounced in Tone 6 as in "si6". She said that for the longest time, she thought that that really was the Cantonese word for "stubborn" and would even use it in spontaneous conversation. It was not until many years later that her sister told her that her step brother was joking around with her. She now knows that the actual word is ngaang5 geng2 (literally "hard neck"). This anecdote illustrates the pervasiveness of correspondence rules. They are so productive that they can even mislead some GEN 2 speakers into thinking a particular word is a Cantonese word.

The regularization of correspondence rules can also lead to the emergence of some lexical forms that are different from what would be used in Hong Kong. This is particularly true for forms that develop pronunciation variants mediated by the writing system (discussed in Section 4.2.3). One example is the word for "apartment". According to Cantodict, the form in use in Hong Kong is [paak3 man4]. One speaker (C2F22A) used the form <aa3 paat1 man3> in the interview. This form is, thus, closer to the Toronto English form in having three syllables. This example also illustrates awareness of "r-dropping" in many GEN 1 English loanwords since both Cantonese and Hong Kong English are non-rhotic. Another example is the word <ci1 si2> ('cheese') introduced in the earlier excerpt from C2F22A. The form recognized by Cantodict is actually <zi1 si2>, a pronunciation influenced by writing. Replacing the initial <z> (an unaspirated alveolar fricative, /t͡s/) with a <c> (an aspirated alveolar affricate, /t͡sʰ/) brings the Cantonese pronunciation closer phonetically to the English form by adding aspiration.

From these examples, we can see that knowledge of Cantonese loan word phonology is part of the linguistic knowledge of GEN 2 speakers. It helps them establish a set of sound correspondences between Cantonese and English allowing them to convert any given English word to a Cantonese word. Using these correspondence rules presupposes knowledge of both Cantonese and English phonology. Thus, the use of correspondence rules contributes to maintaining cross-linguistic phonetic and phonological differences between the two languages. Thus, the lack of vowel shifting among GEN 2 Cantonese speakers is more than simply about awareness of two distinct phonological systems and the social value of speaking "accentless" speech. It is also about putting this knowledge to practice.

To conclude this section, the excerpts presented from interviews with GEN 2 speakers illustrate not only a high level of metalinguistic awareness of cross-linguistic similarities and differences but also how speakers put this awareness into practice. GEN 2 speakers use correspondence rules (Thomason 2007) because they think this is a part of normal Hong Kong Cantonese. GEN 2 speakers also use correspondence rules to fill in gaps in their lexical knowledge of Cantonese. For GEN 2 speakers, the phonology of Cantonese may in fact be the most salient part of the language. This stands in contrast to much of the variationist literature which has described vocabulary as the focus of speaker's metalinguistic awareness of other dialects. For example, Johnstone (2013) has noted that lay person descriptions of Pittsburghese focus on words and phrases rather than on abstract features such as /aw/-monophthongization. Some sociolinguists have even made claims about the inaccessibility of structural features in contact situations (as in Labov's TD Model). That is clearly not the case for GEN 2 Cantonese speakers in Toronto. The use of correspondence rules shows that phonological structure (both segmental and suprasegmental) is accessible for GEN 2 speakers. At the same time, their use could also account for an overall lack of phonetic change.

7.2 ACCOUNTING FOR STRUCTURAL CHANGES

Given the high level of metalinguistic awareness as discussed in the previous section and the lack of inter-generational vowel shifts, it may be surprising that there are some speakers showing evidence of structural changes. The specific changes include the merger of the two high round vowels and splits in /ɛ/. In this section, I argue that these changes are related to the relative frequency with each sound occurs in spontaneous speech. While relative vowel token frequency appears to predict which vowels and vowel contexts are most likely to undergo structural change, individual CAN % and ENG WC Scores predict the specific individual speakers who are most likely to participate in these changes. Thus, the effect of frequency is a probabilistic explanation that accounts for which parts of the vowel system are most susceptible to change rather than which parts of the vowel system will definitely change for all speakers.

Table 67. Token counts for most frequently occurring vowels

Vowel	GEN 1	GEN 2	Intergenerational Changes Observed
/a/	3172	1896	None
/ɔ/	2295	1723	None
/ɐ/	2279	1662	None
/i/	2106	1509	None
/ʊ/	1237	884	None
/ɪ/	1010	625	None

Table 68. Token counts for least frequently occurring vowels (with /ɛ/ separated by context)

Vowel	GEN 1	GEN 2	Inter-generational Changes Observed
/œ/	831	400	Possible F3 changes to be investigated in future study
[ɛ] (open syllable)	703	538	None
/y/	**623**	**351**	**Retraction (Section 6.3)**
/u/	**435**	**165**	**Fronting (Section 6.5)**
[ɛŋ] (pre-nasal)	**342**	**258**	**Fronting (Section 6.4)**
/ɵ/	296	218	Possible F3 changes to be investigated in future study
[ɛ] (stop coda)	**90**	**40**	**Retraction (Section 6.4)**

Table 67 shows a list of the six most common vowels for each Toronto group ordered from most frequent to least frequent. Also included in this table are token counts of each of these vowels for both Toronto groups. The results from the previous chapter showed no evidence of inter-generational change in the F1 and F2 means for any of these vowel categories. Table 68 includes a list of the least common vowels in the data. The token counts for each vowel for both Toronto groups are also included. The vowel /ɛ/ is separated into three different contexts: open syllable, pre-nasal, and stop coda. In contrast to the vowels shown in Table 67, four out of the seven vowels and vowels in phonetic contexts shown in Table 68 do show evidence of change. Two of these vowels (/ɵ/ and /œ/) showed impressionistic variation that would require an F3 analysis to confirm. Only one phonetic context shown in this table showed complete lack of evidence for change: [ɛ] (open syllable). The fact that all of the vowels and vowel contexts that showed evidence for change are among the least common in the corpus suggests that token frequency in spontaneous speech could account for the specific structural changes identified in this study.

Token frequency of different vowel categories is related to language use in conversational context. Table 69 below is a list of the ten most frequent words for the GEN 1 and GEN 2 groups. These rankings are based on raw token counts. Not surprisingly, all of these words are function words rather than content words. Among those that contain a monophthong, the most common is <ŋɔ5> (1^{st} person pronoun). Other words containing a monophthong include <gam2> (‘therefore’), <aa3> (final particle), <go3> (classifier), <di1> (classifier). In each case, the vowel included is among the top four monophthongs according to the list shown in Table 67. The high token frequency of these vowels is thus related to the fact that they occur in many function words, including some that are the most commonly used words in spontaneous speech. Words that appear in both the GEN 1 and GEN 2 top ten list are indicated in bold. The two groups share a total of seven out of the ten top ten words from each group. Thus, it seems likely that this reflects what speakers actually hear in everyday speech. GEN 2 speakers hear and use more tokens of these vowels than tokens of other vowels. This makes it more likely that they maintain the GEN 1 pronunciation of these vowels.

Table 69. Most common words in corpus

GEN 1	GEN 2
Hai6 (‘to be’)	**Hai6 (‘to be’)**
Ngo5 (1st person PRO)	**Ngo5 (1st person PRO)**
Gam2 (therefore)	**Go3 (classifier)**
Go3 (classifier)	**Di1 (classifier)**
Hou2 (intensifier, ‘good’)	Go2 (demonstrative pronoun)
Aa3 (final particle)	M4 (negation marker)
Di1 (classifier)	**Hou2 (intensifier, ‘good’)**
Ge3 (final particle)	**Jau5 (‘to have’)**
Keoi5 (3rd person pronoun)	**Keoi5 (3rd person pronoun)**
Jau5 (‘to have’)	Zau6 (‘and then’)

Token frequency could also be related to phonotactic structure. For instance, the only environment in which /y/ and /u/ contrast is the velar onset environment. As a result of this phonotactic constraint, the number of minimal pairs in which these two vowels contrast is very small. Words that occur in this environment are almost exclusively content words. The only exception is <jyu4 gu2> (‘if’), but neither <jyu4 gyu2> nor <gyu2> are words in Cantonese. Furthermore, <jyu4 gu2> is a variant form of the word [jyu4 gwo2], with the latter being much more common. Since minimal pairs in which /y/ and /u/ contrast are rare in everyday spoken speech, GEN 2 speakers have less exposure to hearing contexts in which this contrast is important. This would make this part of the vowel space more susceptible to change than other parts of the vowel system.

Sound correspondences between Cantonese and English loan words that have entered Cantonese could also account for the specific structural changes observed. As mentioned in Chapter 5, /y/ is the only Cantonese vowel that lacks examples of English loan words (with one debatable exception). A lack of matching English vowels in English loan words with /y/ would mean the lack of a correspondence rule establishing a relationship between Cantonese /y/ and an English vowel. English loan words, however, have entered Cantonese with /u/. This difference between /y/ and /u/ could explain why there was inter-generational shifting of /y/, but not /u/ (as observed in Section 6.3). Even though the retraction of /y/ is part of a merger, this merger has led to more retraction of /y/ than fronting of /u/.

On the other hand, sound correspondences between /ɛ/ in pre-nasal environment in English and in Cantonese, do exist. Cantonese words with /ɛ/ in pre-/ŋ/ environment, however, did not occur as often in the interviews as other vowels. Cantonese /ɛ/ in pre-/n/ environment was even more rare. Thus, for the pre-nasal split in /ɛ/, low token frequency due to the small number of words showing sound correspondences between Toronto English and Hong Kong English could contribute to making this part of the vowel system susceptible to change. There are also English loan words that have entered Cantonese with /ɛ/ in other environments. Many of these loan words likely entered Cantonese through UK English during the colonial period of Hong Kong's history. The English source words vary in corresponding to either /ɛ/ or /æ/ in Toronto English. Examples of these words include [fɛn55] ('friend'), [pœ22 sɛn55] ('percent'), [ɛn55 t͡sin35] ('engine'), [pɛn55] ('band'), [kʰɛn55 sa35] ('cancer'), and [mɛk22 kʰɛn55] ('mechanical'). It could be the lack of clear and consistent cross-linguistic sound correspondences that explain why this is a part of the vowel space that has become susceptible to change among some GEN 2 speakers. In this case, however, it is not about loss of phonological contrast. It is about increasing phonological complexity influenced by cross-linguistic similarity with this part of the vowel system.

The fact that there is the structural potential for English to affect Cantonese does not mean that all GEN 2 speakers are affected. Instead, there is individual speaker variation. Phonetic similarity and token frequency matter only in terms of accounting for the likelihood of which parts of the vowel system will undergo cross-linguistic influence. Factors identified in Chapter 6 that best account for why some speakers have been influenced by Toronto English and why some have not include CAN % Scores and ENG WC Scores. These factors suggest that those who used less Cantonese and more English vocabulary in their interviews were the ones to lead in these changes. These factors point to linguistic choices.

Aside from CAN % Score and ENG WC Score, only one other factor came out significant in accounting for both the retraction of /y/ and the fronting of /u/: EOQ Question C1 ("What language does your family speak when you get together?"). Those who speak with their families in Cantonese are more likely to keep the two vowels distinct than those who speak with their family members primarily in English. This further supports the argument that language use is a primary factor in accounting for innovation in this part of the vowel system. The results presented in Chapter 6 also showed some factors predicting /y/ retraction but not /u/ fronting. These include EOQ A1 ("Do you think of yourself as Chinese, Canadian, or Chinese-Canadian?") and EOQ A5 ("when you were growing up, were the kids in your school Chinese?"). Both of these questions relate to individual speaker identity. The fact that the responses to these two questions came out significant in the retraction of /y/ suggests that identity can partly account for the incipient merger of /y/ and /u/. Identity itself, however, may be directly related to language use. Those who identify as more Chinese rather than Canadian are also likely to be the ones who are able to use Cantonese across a wider range of contexts. Language use, thus, appears to be important in ensuring GEN 2 speaker ability to produce a phonological contrast in the two high round vowels.

Many factors that have been shown to account for sound change in the variationist research literature did not come out significant for any of the structural changes observed in the current study. This includes gender. "Sex" was not a significant predictor of /y/ retraction, /u/ fronting, or fronting of pre-nasal /ɛ/. Age also did not come out significant suggesting that these structural changes are more directly tied to contact with Toronto English than to the emergence and spread of an apparent time change. None of the EOQ Factors other than the ones mentioned above for /y/ and /u/ came out significant either.

For the pre-nasal split in /ɛ/, the only significant factors were CAN % Score and ENG WC Score. One possible explanation for the lack of significant predictors for the /ɛ/ nasal split compared to the merger of /y/ and /u/ is the different nature of the structural change involved. While a merger leads to a loss in phonological contrast, a split increases phonological complexity. A split could have a broader effect across all GEN 2 speakers since it does not result in loss of meaning. This could mean that fewer Cantonese speakers notice the split. For example, in Section 6.4.1, I presented an excerpt from an interview in which one speaker mentions how some people have noted her unusual pronunciation of the vowels involved with the /y/~/u/ merger. There was not a single case, however, of anyone mentioning anything about the pronunciation of words with /ɛ/ in pre-nasal context. Although this is worth further investigation, it could be the case that while the /y/~/u/ merger is a change above the level of conscious awareness, the pre-nasal /ɛ/ split is below the level of consciousness. This could explain the lack of significant predictors other than CAN % and ENG WC scores in accounting for the pre-nasal

/ɛ/ split. Thus, this is still a structural change related to frequency of use. With exposure to fewer examples of pre-nasal /ɛ/ in spontaneous speech, GEN 2 speakers would be more likely to be influenced by English phonology in their production of vowels in this part of the Cantonese vowel space. At the same, a split may also be less salient to other Cantonese speakers since a split is not an innovation that leads to loss in phonological contrast.

To conclude this subsection, relative vowel token frequency appears to predict which vowels and vowel contexts are most likely to undergo structural change while individual linguistic choices (based on CAN % and ENG WC Scores) appear to predict the individual speakers who are most likely to participate in these changes. Thus, the effect of frequency is a probabilistic explanation that accounts for which parts of the vowel system are most susceptible to change rather than which parts of the vowel system will definitely change for all speakers.

7.3 PROBLEMS WITH THE TD MODEL

As discussed in Chapter 2, the TD Model is a model of sound change rooted in the basis between the difference between child and adult language acquisition. It draws parallels between these two acquisition processes and the socio-historical outcomes of sound change. Child language acquisition is described as "regular" and leading to Neogrammarian sound change while adult language acquisition is formulated as introducing lexical irregularities into a system leading to patterns of change distinct from those initiated by children. The findings from the current study present challenges to the dichotomy presented in this model. They show that changes initiated by

GEN 2 Toronto Heritage Cantonese speakers do not completely fit the characteristics described as either "transmission" or as "diffusion".

One finding from this study is the lack of overall GEN 2 initiated changes. In this respect, there is almost "perfect" transmission of Cantonese phonology from GEN 1 speakers to GEN 2 speakers. This "perfect" transmission also applies to low-level phonetic patterns such as higher F1 in closed syllable context for /ɔ/. GEN 2 speakers, also, seem to participate in at least one change in progress developed in Hong Kong (and also present in GEN 1 speech). This change is the apparent time fronting of /i/. GEN 2 speakers, however, have not advanced this change as far as Homeland speakers have. Thus, GEN 2 appear to have slowed down or stalled this change. In one sense, GEN 2 speakers appear to display more "perfect" transmission than Hong Kong speakers. The only other vowel shift identified appears to be part of an incipient merger rather than part of a vowel chain shift. The Homeland group also shows more examples of vowel shifting overall. They all appear to be consistent with Labov's (1994) Principles of Vowel Chain Shifts.

The other major finding is that the only GEN 2 innovations are a merger and splits. The structural aspects of these changes are consistent with Transmission, but what is not consistent is the finding that these are contact-induced changes. The evidence for contact-induced change comes from the lack of the same changes in the Homeland data and results showing that those who lead in these innovations are the ones with lower CAN % Scores and higher ENG WC Scores. Labov argues that structural changes cannot be introduced by diffusion (contact) except for mergers as in Herold's (1990; 1997) study of the low-back merger in Northeastern Pennsylvania. His explanation is that it is easier to lose a distinction than to gain one. Thus, gaining a new distinction as in the case of learning a phonemic or allophonic split is something that can only be done in child language acquisition and hence through transmission. Labov sums up this point by stating that "the distinction between transmission and diffusion is maximal in the case of splits" (2007:371). The structural changes identified among GEN 2 speakers, however, include not only a merger but also vowel splits. All of these changes are contact-induced and hence through diffusion rather than transmission.

The other challenge that the findings present is that these contact-induced changes affect Cantonese inherited vocabulary with "regularity". This study focused exclusively on Cantonese vocabulary. Although some English loan words that have become well integrated into Cantonese were included, none of these English loan words were involved in these changes. For instance, as discussed in Chapter 5, /y/ is the only vowel for which there is a complete lack of English loan words (with only one disputable exception). All the words involved in the merger identified among the speakers that have them are part of Cantonese native vocabulary. Borrowed words with /u/ (pronounced in both Toronto English phonology and in Cantonese phonology) were intentionally excluded. For the nasal split in /ɛ/, the nasal class words are also exclusively from Cantonese native vocabulary, yet they are all affected by the change. Thus, although these innovative pronunciations are likely influenced by contact with English (thus structural diffusion), they show more Neogrammarian like patterns (transmission). They do not show the "irregularity" and randomness that Labov has described as characteristic of contact-induced change (and in particular lexical diffusion).

As discussed in Chapter 2, the TD model rests on key assumptions that are not applicable in this particular context. Labov specifically says that his model "is dependent on the concept of a speech community with well-defined limits, a common structural base, and a unified set of sociolinguistics norms" (2007:347). The results from this study suggest two other implicit assumptions that are not applicable in the context of Toronto Heritage Cantonese. One is that individuals can have only one "natural" phonology. Some GEN 2 speakers show evidence that they have at least two distinct phonological systems. For example, in Section 7.1, I discussed how C2M21C is able to produce an F1 distinction between Toronto English [ɑ] and Cantonese [ɔ]. The second implicit assumption is that the vernacular (and L1 phonology in general) is immutable once it has been acquired. All GEN 2 speakers are L1 speakers of Cantonese. Over time, they all learned English as a second language. Most (if not all) GEN 2 speakers are English dominant as evidenced in their near unanimous responses to some of the EOQ questions related to language use discussed in Section 5.3.2. All but one speaker said that their preferred language is English rather than Cantonese. Similarly, all but one speaker said that that the language they use when talking to friends is English. This has made it possible for some GEN 2 speakers to have their English phonology influencing their Cantonese phonology. In other words, it is the L2 that is now the dominant language for GEN 2 speakers and the L2 for some speakers is now affecting their L1 (vernacular). The vernacular for these GEN 2 speakers is, thus, not as immutable as assumed by the TD Model.

To sum up this section, I have discussed how the results from this dissertation study are inconsistent with the strict dichotomy between transmission and diffusion proposed by Labov (2007). Variation and change among GEN 2 Toronto Heritage Cantonese speakers show characteristics of both Transmission and Diffusion. For example, all GEN 2 speakers learned Cantonese as their L1. This, Labov argues, is the basis for Transmission. Some of the patterns observed are characteristic of Transmission such as a lack of overall intergenerational change and the initiation of structural changes that occur with Neogrammarian regularity. On the other hand, I have also shown evidence that these structural changes have been influenced by a second language. Labov describes second language acquisition as the basis for Diffusion. Thus, GEN 2 Toronto Heritage Cantonese speakers show evidence of both Transmission and Diffusion, which does not seem consistent with the assumption of a strict dichotomy between these two processes. I have also discussed how some of the key assumptions behind the TD Model are not assumptions that apply to GEN 2 Toronto Heritage Cantonese speakers. For example, the TD Model implicitly assumes that speakers can have only one phonology and that the vernacular (which is treated as synonymous with the first acquired language) is immutable. Some of the speakers analyzed, however, show evidence for two distinct phonological systems. Results showing that some speakers can be influenced by the phonology of a second acquired language challenge the assumption of a phonologically immutable vernacular.

7.4 THE THOMASON & KAUFMAN VS VAN COETSEM MODELS

The results from the current study do not pose the same problems for either the TK model or the VC model (discussed in Chapter 3) as they do for the TD Model. Both the TK and the VC models recognize genetic and typological similarity as facilitating factors in the outcomes of contact-induced change. They also both recognize the possibility of structural influence in communities with early bilingualism. Where these two models differ is in how they would describe structural influence in cases involving early bilingualism. The results from the current study have implications for addressing this issue.

Both the TK and VC models are similar in recognizing two general types of transfer in contact-induced change. The first is "borrowing" while the second is called either "shift-induced interference" (TK model) or "imposition" (VC model). Under the TK model, "borrowing" has a much broader meaning than under the VC model. "Borrowing" is tied to language maintenance and socio-historical continuity in the language transmission process. Under the TK model, Toronto English influenced changes such as the merger of /y/ and /u/ and the pre-nasal split in Cantonese are cases of structural borrowing. Cantonese continues to be transmitted to a new generation of speakers. The genetic heritage from Proto-Cantonese (and Proto-Chinese) is maintained. While the TK Model seems adequate for describing broad community-level changes from a socio-historical perspective, what is lacking is specific discussion of individual speaker mechanisms that give rise to change and how these changes are propagated across a community of speakers.

The VC Model attempts to address the question of the individual by basing a distinction between the two transfer types on individual speaker linguistic dominance rather than on community level language dominance. These two mechanisms are described as "neutralized" in cases in which speakers are fluent in two languages. This means that for balanced bilinguals, it is difficult to distinguish between borrowing and imposition. For the conservative GEN 2 speakers, this distinction does not matter since they show lack of evidence for change. The innovative speakers, however, are the ones who had lower CAN % Scores and higher ENG WC Scores. This would suggest that for this group of speakers, the mechanism behind change is imposition according to the VC Model. In fact, one argument that has been raised against the TK model is the failure to distinguish between individual level and societal level linguistic dominance. Some speakers become linguistically dominant in their L2 and once this happens, structural transfer would be influenced by the L2 rather than by the L1. Critics of the TK Model (see Smits 1998;

van Coetsem 1990; Winford 2003; 2007) have thus argued that a heritage language contact case would be better described as imposition rather than borrowing.

The findings from the current study, however, provide only partial support to the assumption that phonological change among heritage speakers is identical to imposition introduced by adult L2 speakers. While the lack of an /y/ vs. /u/ contrast and a pre-nasal /ɛ/ split could certainly be possible for an L1 Toronto English speaker learning Cantonese as an adult (at least in initial stages of learning), this dissertation does not address the speech of such speakers. Thus, it is not known for sure whether or not such speakers would have these features in their Cantonese speech. Furthermore, this dissertation addresses only monophthongs. The current study does not consider whether or not Toronto English has influence on other phonetic or phonological aspects of GEN 2 Cantonese speech such as suprasegmental features, diphthongs, or consonants. Experimental studies that do address phonetic production among both heritage speakers and adult L2 speakers of the same language, however, suggest that these two groups of speakers are phonetically different from each other.

For example, one study discussed in Section 3.3 was Chang et al.'s (2011) study comparing heritage speakers of Mandarin with adult L2 speakers of Mandarin. Both groups speak English as their dominant language. This study found that the /y/ vs. /u/ contrast in Mandarin is implemented differently in these two groups. The L2 Mandarin group appeared to be more influenced by their knowledge of English than did the heritage speaker group. The results, thus, show that the two groups are phonetically distinct from each other. Imposition (according to the VC Model) does not lead to the same outcome in heritage speakers compared to adult L2 speakers according to the results of this study. Models of contact-induced sound change, thus, need the capacity to be able to account for this difference. The VC model assumes that the same psycholinguistic process applies to both groups. This may be true, but if it is true, the process applies to individuals with different language developmental histories. These earlier developmental histories would affect the different outcomes observed in the two groups.

The results of this dissertation study also suggest a need to distinguish between linguistic dominance and proficiency. The VC Model treats these two concepts as synonymous. While all the GEN 2 speakers analyzed speak English as their dominant language, there is clear variation in their proficiency levels (at least in terms of their CAN % and ENG WC Scores). The evidence for linguistic dominance in English comes from responses to various EOQ questions. All but one speaker said that their preferred language and the language they would use with friends is English. If most of these speakers are linguistically dominant in English, we would expect these speakers to show influence from English in their Cantonese speech. Only a subset of these speakers, however, show evidence of merger of /y/ and /u/ and a pre-nasal /ɛ/ split. Specifically, it was speakers who have lower CAN % and higher ENG WC Scores that lead in these changes. This raises the question of what exactly these scores measure.

The CAN % Score measures the amount of Cantonese used in the recorded interview sample for each speaker while the ENG WC measures the total number of unique English words. Speakers could have various reasons for using less Cantonese or more English. On the topic of code-switching rates, Lyskawa et al (2016) say that code-switching could either be something that "demonstrates a native-like command of both languages" or "an avoidance strategy for navigating around the most complex structures or vocabulary in a language in which a speaker is not fully confident" (Lyskawa et al 2016:227). Similarly, the relative amount of Cantonese and spoken in these interviews as well as the number of unique English words uttered could either be because speakers are highly fluent in both languages and can code-switch or because speakers have weak proficiency in Cantonese.

There is evidence that for at least some of these speakers, low CAN % and high ENG WC scores are an indication of weak Cantonese proficiency. C2F24A, for example, has both the lowest CAN % Score (29.36%) and the highest ENG WC Score (1003). She, thus, used the least amount of Cantonese and the most amount of English among the speakers analyzed. An excerpt from her picture naming task was presented in Section 6.4.1. The excerpt is almost entirely in English syntax. The Cantonese used is limited to lexical items. We can see C2F24A asking the interviewee, “Did I say it wrong?” C2F24A does not show very strong confidence in her Cantonese speaking abilities. Although there were some parts of her interview in which she produced complete sentences in Cantonese, it is clear that she struggled more than any other speaker in speaking Cantonese. Thus, for C2F24A, it seems clear that her low CAN % Score and high ENG WC scores are due to weak proficiency in Cantonese. Her relatively high use of English was primarily to address gaps in her Cantonese speaking abilities.

For other speakers analyzed, however, the extent to which CAN % and ENG WC Scores are an indication of relative proficiency levels is uncertain. The reasons for using English in an interview designed to be primarily in Cantonese appear to be varied. High proficiency speakers could have decided to use more English because those speakers encountered more contexts in which English seemed more appropriate to them. The CAN % Scores for these speakers may be artificially lower than expected. Some lower proficiency speakers may also have CAN % Scores that are higher than expected if they fill in lexical gaps in their Cantonese knowledge through the use of correspondence rules (as discussed in Section 7.1). Without conducting a closer analysis of code-switching patterns, it is uncertain whether or not there is a predominant reason for code-switching. In any case, there does seem to be evidence that for at least some speakers, CAN % Score and ENG WC Score are related to proficiency. If this is true, we can see an important need to distinguish between linguistic dominance and proficiency, a distinction not made in the VC Model. Whatever is the reason for code-switching, the results from this study appear to be consistent with Lyskawa et al in showing that "code-switching provides the context in which the rules of both languages are active and employed in a convergent manner" (2016: 236).

Smits (1998) described one of the key differences between the two models as one that is focused on a socio-cultural (TK) vs. a psycholinguistic (VC) perspective. In fact, critics of the TK Model assume that a model focusing more on the psycholinguistic mechanisms of change (as in the VC Model) would provide a more accurate picture about the actuation of change. The problem with this line of reasoning is that there is still a lot that researchers are learning about heritage speaker phonetics and phonology from both a psycholinguistic perspective and from a socio-cultural perspective. Since the publication of both the TK and the VC Models, there has been a growing body of research showing how heritage speakers may be phonetically different from both prototypical L1 and adult L2 speakers of the same language. A psycholinguistic perspective does not necessarily provide the basis for a better model if it is based on inaccurate assumptions.

Critics of the TK Model, however, have raised a valid point about the difficulty of interpreting changes initiated by heritage speakers and others who become linguistically dominant in their L2. This only necessitates the need for more research on heritage speakers especially from a variationist perspective, which is lacking in the current body of literature on heritage languages. By analyzing heritage speaker phonetics and phonology based on conversational speech data, the current study presents a perspective that differs from both controlled experimental settings and classroom settings. Results showed phonological variation, which challenges the claim made by Polinksy and Kagan that "one could easily imagine that there would be no differences in the phonological representations between the heritage language and the baseline" (2007:378). This was made possible using the methods for the study of sound change in progress pioneered by Labov, Yaeger, and Steiner (1972). Conversational speech data also yielded valuable sociolinguistic insights into how GEN 2 speakers actually use Cantonese. The use of correspondence rules as well as metalinguistic discussions about them, for example, showed how important they are in the maintenance of Homeland phonology. These findings demonstrate that it is ultimately how speakers use their language and the social meanings they construct about language use that shape the direction of change. Such insights could easily be missed in controlled experimental studies since the messiness of language mixing is often a factor to be avoided than to be studied. If there are any cognitive or articulatory constraints in this process, they are clearly weaker and fewer than what has been observed in the majority of research conducted on monolingual speech communities.

What may be of crucial importance from a sociolinguistic research perspective is developing a better understanding behind the social meaning of phonological variation and the use of different types of codes (ex: English words pronounced with Toronto English vs. Cantonese phonology) within a community of heritage speakers. Assuming Cantonese continues to be passed down to future generations of speakers, do social norms and practices favor maintenance of more conservative or more innovative phonology? Although it is clear from the results from this study that structural change is possible on the individual level, there is also evidence from the interviews that more conservative forms are favored. If this is the case, will the GEN 2 innovations identified in this study be passed down to future generations of speakers?

This is an open question that cannot be answered based on the results from this study. What this study has been able to demonstrate is that sound change (or even the lack of it) within a community of heritage speakers is a worthwhile topic for future research.

7.5 TOWARDS AN IMPROVED MODEL OF SOUND CHANGE

As mentioned in the previous section, one of the main differences between the TK and the VC Models is that the former takes a socio-historical perspective while the latter takes a psycholinguistic perspective. In this section, I discuss what needs to be part of an ideal model of contact-induced sound change. I argue that an ideal model would be one that integrates multiple perspectives. I will specifically discuss four different types of factors and how such factors were needed to account for the specific findings of the current study. These factors include macro-level social factors, micro-level social factors, linguistic factors, and individual speaker factors.

7.5.1 Macro-level Social Factors

The first type of factor I discuss are macro-level social factors. By macro-level, I include factors related to large demographic groups, social organization, socio-economic factors, and socio-political factors. The findings from the current study suggest that macro-level social factors are not deterministic. Yet, they are important in establishing the setting under which contact-induced change can take place.

First of all, immigration from Hong Kong to Toronto is what has made Toronto English influence on Cantonese possible. As discussed in Section 4.3, the reasons for immigration included both socio-economic and socio-political reasons. Changes in immigration laws in Canada, for example, were important in making it possible for large numbers of Hong Kong immigrants coming to Toronto. Also discussed in Section 4.3 was how some of these laws favored more affluent immigrants. This made it possible for many of these immigrants to purchase property in Toronto and to create businesses catering to the local Cantonese-speaking community. This would subsequently pave the way for even more Cantonese-speaking immigrants to Toronto, with latter waves being more socio-economically diverse than earlier waves. From the 1980s through the 1990s, one of the major reasons for immigration to Canada was fear of what would happen to Hong Kong after the 1997 handover back to China. In short, socio-economic and socio-political factors are important because they account for why Cantonese speakers came in contact with Toronto English speakers.

The local dominance of English in Toronto is also important to consider. What this means is that most GEN 2 speakers born and raised in Toronto become English-dominant speakers of Cantonese. While there may be some individuals who are exceptions, the specific set of speakers analyzed unanimously identified English as their preferred language as well as the language they most often used when speaking with friends. In addition, every single speaker analyzed said that they preferred to read and write in English. The local social dominance of English in Toronto, thus, accounts for why most (if not all) GEN 2 speakers are also linguistically dominant in English.

7.5.2 Micro-level Social Factors

Macro-level social factors alone cannot tell the complete story. The role they play is merely in establishing the setting under which possible contact-induced changes can occur. For example, as discussed above, we know that English is the locally dominant language in Toronto. We also know that Cantonese is also spoken by many speakers in Toronto because of mass immigration from Hong Kong. We also know that there are Cantonese can be used as a language in service encounters because of some of these immigrants established local businesses that cater primarily towards Cantonese-speaking customers. Micro-level social factors relate to how speakers go about their day-to-day lives in such a setting. They also show that macro-level social factors are not deterministic because sometimes speakers can resist forces that have otherwise been set in motion by macro-level social forces.

For example, as discussed in Section 7.1, GEN 2 speakers recognize that English and Cantonese have different phonological systems. They are also able to produce distinctions between Cantonese vowels and phonetically similar Toronto English counterparts when it becomes socially meaningful to produce a distinction. This was illustrated through discussion of one speaker's different pronunciation of the name of "Loblaw's", the name of a local supermarket chain. Also discussed in Section 7.1 were examples of how this knowledge is linguistically productive. GEN 2 speakers recognize sound correspondences between Toronto English and English loan words in Cantonese (which generally corresponds to the non-native English spoken by GEN 1 speakers). Sometimes they overextend their knowledge to create new words as was shown in a few interview excerpts. I also presented examples in which GEN 2 speakers filled in lexical gaps in their knowledge of Cantonese by using English with Cantonese phonology. This shows how speakers put their awareness of two distinct phonological systems into practice.

How speakers actually use their languages is important because speakers have choices in how to use their languages. There may be some bilingual communities in which borrowed vocabulary is pronounced with source language phonology intact. For GEN 2 Toronto Cantonese speakers, this is sometimes the case, but the examples discussed in Section 7.1 show that pronouncing English vocabulary using Cantonese phonology is also possible. The fact that this is possible illustrates how knowledge of two distinct phonological systems is productive knowledge. The meaningful use of this knowledge could partly account for why there is an overall lack of contact-induced vowel shifting in GEN 2 Cantonese speech. Thus, micro-level social factors, such as metalinguistic practices, must be considered along with the macro-level social forces that make such micro-level metalinguistic practices possible.

7.5.3 Linguistic Factors

A third important consideration is linguistic factors, both in terms of structural properties and in terms of frequency of occurrence in spontaneous speech. As discussed in Section 7.2, such linguistic factors appear to account for which parts of the Cantonese vowel system are most susceptible to change.

For example, one part of the system that seems to be susceptible to change is the high round tense vowels: /y/ and /u/. Based on token frequency information, I showed that /y/ and /u/ are both among the least frequently occurring vowels in the interview samples. This is the case across all three speaker groups. Frequency could also be related to structural factors. /y/ and /u/

are in near complementary distribution in Cantonese. They are contrastive only in velar onset context. Thus, not surprisingly, the number of actual minimal pairs is relatively small. Most of the speakers analyzed did not have a single minimal pair. The fact that the contrast between /y/ and /u/ is not a robust contrast in Cantonese could account for why this is a part of the vowel system that is susceptible to contact-induced change.

On the other end of the token frequency scale are vowels such as /ɔ/. There is no evidence for inter-generational lowering of this vowel towards the phonetically similar Toronto English /ɑ/. This may be due to the fact that /ɔ/ is among the most frequent vowels in Cantonese spontaneous speech. In fact, this vowel occurs in many functions words such as the first person pronoun, [ŋɔ23], and the generic classifier, [kɔ33]. Thus the fact that there is lack of lowering of /ɔ/ but merger of /y/ and /u/, at least among some speakers, shows that frequency of occurrence is an important factor in accounting for susceptibility to change.

The other part of the vowel system that shows susceptibility to change is /ɛ/ in pre-nasal context. This is a part of the vowel system that corresponds to two phonetically different allophones in Toronto English. As I discussed in Section 7.2, /ɛ/ in pre-nasal context is also in the lower half of vowels in terms of token frequency. I also discussed how there is a lack of consistent sound correspondences between the pronunciation of Cantonese words borrowed from English and the pronunciation of these words in Toronto English. In some words, Cantonese [ɛ] corresponds to Toronto English [ɛ] as in [fɛn55] ('friend') but in other loan words, Cantonese [ɛ] corresponds to Toronto English [æ] as in [bɛn55] ('band'). The lack of consistent sound correspondences in this part of the vowel space based on loan words could account for why this part of the vowel space shows susceptibility to change. Such linguistic factors are, thus, also important in understanding contact-induced sound change.

7.5.4 The Individual

Finally, the last type of factor I discuss is individual speaker factors. As discussed above, macro-level social factors set the contact setting. Micro-level social factors relate to the social meaning behind linguistic distinctions in a given contact setting. Linguistic factors such as frequency and structure condition the parts of the phonological system that are most susceptible to change. While these three types of factors all seem to condition and constrain the possibilities in terms of contact-induced change, whether or not contact-induced change actually happens appears to be individually-based.

Many GEN 2 speakers show little or no evidence of contact-induced change. This includes lack of vowel shifts, mergers, and splits. The speakers that do show contact-induced changes, however, do seem to have something in common. Specifically, it was lower CAN % and higher ENG WC Scores that predicted which individual speakers were the most likely to have merger of /y/ and /u/ and a pre-nasal /ɛ/ split. Language used at home also accounted for the merger of /y/ and /u/ but not for the pre-nasal /ɛ/ split. All of these factors point to individual language use patterns. As discussed in Section 7.4, there is evidence that the speakers with the lowest CAN % and highest ENG WC Scores have weaker proficiency skills than other speakers. Thus, consistent with the VC Model, proficiency appears to play a role in determining which specific speakers are the ones most likely to initiate structural changes.

The role of proficiency highlights the role of the individual speaker. At the same time, it seems unclear how innovative heritage speakers propagate change across the community. There is evidence showing that conservative forms are valued in this community. If more proficient speakers are more likely to pass down Cantonese to subsequent generations of speakers, these more proficient speakers would pass down more conservative forms since these are not the speakers who are innovating. If the innovative speakers are the least proficient, it is unclear to what extent their pronunciation patterns may be passed down to subsequent generations of speakers. In any case, what the results of this study have been able to demonstrate is that the role of individual speaker factors needs to be considered along with other factors.

7.6 FUTURE RESEARCH DIRECTIONS

I conclude this chapter by discussing possible research directions to further support the claims made in this chapter. The current study has several limitations with respect to acoustic data. One of them is that it did not include acoustic data on implementation of tonal contrasts. This is relevant because the word tokens pronounced with innovative vowel variants did not appear to be innovative in terms of tonal contour patterns. They were impressionistically the same as GEN 1 patterns. Impressionistic observations are not always accurate. Nevertheless, it does seem likely that if there is tonal merger or complete loss of tone, such changes would have resulted in difficulty transcribing these utterances. This was not the case. Assuming that this observation is correct, it is only segmental structure that has been influenced by Toronto English. If this is imposition, it appears to be only partial rather than complete imposition.

Another limitation of the current study is that it does not include acoustic data from the adult L2 Cantonese spoken by L1 speakers of Toronto English. Such speakers, however, are extremely rare. For the most part, Cantonese exists primarily as an in-group language that is rarely learned by those who do not have Cantonese heritage. Yet, this is an important point to address with respect to the question of imposition because in many cases of change initiated by imposition, the agents of change are those who learn the target language as adults. In the Toronto Cantonese case, however, such a group of speakers is extremely small. It is doubtful that such a group could have an impact on the Cantonese spoken by those with Cantonese speaking parents. Instead, imposition (following the VC Model) would be initiated by L1 Cantonese speakers who subsequently become dominant in Toronto English.

The critical question to ask with respect to the issue of partial imposition is if a native speaker of Toronto English who learns Cantonese as an adult would pronounce words like <gun3> in the same way as one of the GEN 2 speakers who has a more innovative system. Since Toronto English lacks tone, we would expect imposition to involve not only segmental transfer, but also suprasegmental transfer. Such a speaker, at least in the initial stages of learning, would be expected to pronounce <gun3> with a fronted /u/ and perhaps primary stress rather than with a mid-level tone. This is different from what has been observed among innovative GEN 2 speakers. For this group, only the vowel pronunciation has been influenced by Toronto English. The tone categories remain intact. Thus, even if we assume that the same psycholinguistic process is operating on both groups of speakers (as in the VC model), the outcomes are different.

Finally, the third empirical limitation of the current study is lack of data from the English spoken by GEN 2 speakers. This could provide stronger evidence of distinct outcomes between the GEN 2 speakers and hypothetical adult L2 speakers of Cantonese. Chang et al.'s (2011) study also provides support for this possibility. The results from this study showed that heritage Mandarin speakers produced tokens of /y/ at a midway point between the /y/ produced by non-heritage L1 Mandarin speakers and adult L2 Mandarin speakers. This suggests that even though there may be English influence on heritage speaker phonetics, the influence means approaching the acoustic range of phonetically similar vowels without complete cross-linguistic merger. In other words, it could be true that the pronunciation of /y/ among more innovative GEN 2 speakers has approached the pronunciation of Toronto English /u/ while remaining phonetically distinct. A hypothetical group of adult L2 speakers of Cantonese, however, may be more likely to cross-linguistically merge Cantonese /y/ with Toronto English /u/. If this proves to be the case with additional acoustic data, then there would be stronger evidence that imposition (according to the VC Model) is not the same for speakers who acquired different languages as their L1.

8.0 CONCLUSION

I conclude this dissertation by returning to Labov's famous quote presented in the first chapter. Labov "resist[ed] the term sociolinguistics … since it implies that there can be a successful linguistic theory or practice which is not social" (Labov 1972:xiii)? This dissertation is first and foremost a ***linguistics*** dissertation. It is specifically a contribution to the vast body of sociolinguistics research on vowel variation with the aim of studying sound change in progress. As was the case for Labov's work, the Uniformitarian Principle also motivated the current study but where it diverges from earlier studies is the attempt to address sound change in communities, both past and present, characterized by speakers who have acquired two languages at an early age. My approach involved focusing on the vowel system of a language that has not been previously studied through a variationist lens and on a language spoken in a heritage language context and thus potentially involving L2 to L1 influence. Thus, by pushing variationist sociolinguistics research **beyond its monolingually oriented core** and by studying a heritage language **out in the wild** and thus outside a controlled context, my work is a contribution not only to sociolinguistics but to the relatively new area of research on heritage speaker phonetics and phonology.

In Section 8.1, I discuss how my dissertation contributes to ongoing debates within sociolinguistics about the nature of the "social" part of "sociolinguistics". The analysis of my results supports the third wave variationist emphasis on micro-level social interaction (or lack of) and on the individual as primary forces in driving sound change (or lack of change). Another major research finding is that not all heritage speakers have the same phonology. This challenges recent research findings on heritage speakers in controlled settings. I discuss the implications of the variability observed in Section 8.2. Finally, in Section 8.3 I conclude this final chapter by discussing the relevance of the Uniformitarian Principle and how the results of this dissertation can help clarify the phonological development of languages spoken under intense contact settings.

8.1 BEYOND THE MONOLINGUAL CORE

Chapter 2 of this dissertation focused on the history of the different dichotomies that have developed in the study of sound change. Labov's quote presented above reflects a tension long present in linguistics between 'language' and 'everything else'. This dissertation contributes to this ongoing discussion about linguistic structure and the various social forces that can have an effect on it.

First, I point out the irony in Labov's quote in light of critiques of his framework for studying sound change. Coupland, for example, has described the Labovian paradigm as one "in which (for cogent empirical reasons) a highly restricted, featural conception of 'language' is fused with a rather asocial conception of both 'language' and 'change'" (2014:280). This sentiment is shared among other researchers who adopt discursive and third wave approaches. This makes Labov's quote seem ironic since he was clearly trying to push the view that language cannot be isolated from social factors. The problem was his attempt to integrate a theory of change in linguistic structure with a theory of social structure influenced by 1960's era sociological theory. Sociology, however, has since advanced far ahead of the ideas developed in the discipline in the 1960s. So has anthropology, psychology, and other fields related to linguistics. Recent critiques of the Labovian framework have only continued the discussion between 'language' and the social forces that can have an effect on it. Some researchers have pushed even farther in emphasizing the inseparability of language and the social world than Labov ever has.

Pushing variationist research beyond its monolingual core (Nagy 2016), as indicated in the title of this dissertation, is one of the key contributions of my dissertation to this ongoing discussion. Following Thomason and Kaufman's (1988) framework as an alternative to Labov's (2007) TD Model, I have emphasized that the possible outcomes of sound change in a language maintenance under intense contact setting are different from what has been observed in the bulk of sociolinguistic settings focused on monolingual communities. I have illustrated how variation and change in a heritage language can provide an example of the linguistic outcomes of contact between two genetically and typologically distinct languages. The problem presented in heritage language variation and change is that it boils down to individuals acquiring two phonological systems as children. This is not a possibility recognized in Labov's TD Model nor has it been a major emphasis in most variationist research outside of the HLVC Project. It becomes clear that the state of the art in theorizing about sound change has failed to consider all possible types of contact settings as well as all possible directions of influence. I have, thus, addressed one of the research gaps identified by Thomason and Kaufman when they say that "what is needed is research on current or recent contact situations that permit a more ambitious analysis of sociolinguistic context than we have attempted here" (1988:213). Even after 30 years, this remains a major research gap.

An underlying assumption in this dissertation is that the language part of the dichotomy matters because the availability of genetically and structurally distinct phonological systems can have an effect on the dynamics of sound change. I have argued that the reason my results are unusual for variationist studies is because I am dealing with a heritage language spoken in a multilingual community rather than a language spoken as the societally dominant language in which monolingualism is the norm. This is why models of contact-induced change are important because they implicitly recognize a distinction between language contact and dialect contact unlike Labov's TD Model. This assumption goes against Labov's claim that the difference between language contact and dialect contact does not matter since the dialect vs. language distinction is not a linguistic issue (Labov 2007:347 FN1). Thus, by showing the possibility of contact-induced structural influence in Toronto Heritage Cantonese, I have shown problems with Labov's assumption.

The linguistic resources available in a particular sociolinguistic setting matter in accounting for sound change (or lack of it) because these resources set the internal ecology (Mufwene 2001) from which specific changes initiate. A multilingual community has a larger linguistic feature pool (Mufwene 2001) than a multidialectal community because more linguistic resources are available in a multilingual community. Structural similarity across the different languages has an effect on the dynamics of change since similar structures are both more easily conflatable (for both analysts and speakers) and more likely to persist. Other researchers have also supported the assumption that I make that the dynamics of contact-induced change can be different depending on the genetic and typological similarity between the languages involved (see Epps, Huehnergard, and Pat-El 2013; Law 2013; Mithun 2013 for morphology). Thus, the possible outcomes of change in Cantonese among speakers who also speak English are different from the possible outcomes of change that arise in contact between speakers of St. Louis and Chicago English. This is why the TD Model can work in accounting for how the Northern Cities Chain Shift spreads from Chicago English to St. Louis English through lexical diffusion rather than through direct structural diffusion, but it does not work as well in accounting for direct contact-induced structural change among innovative Toronto Heritage Cantonese speakers. Cantonese in contact with English involves far less overlap in similar form-to-meaning mappings than St. Louis English in contact with Chicago English. Furthermore, the Toronto case also involves early acquisition of two distinct languages rather than multiple dialects of the same language.

To be clear the importance of the feature pool (Mufwene 2001) does not diminish the role of social factors. The specific contents of the feature pool are socially transmitted from one speaker to another and from one generation to the next. The importance of the feature pool is merely in setting the starting point from which specific changes arise. Yet, it is important to recognize how the dynamics of change can be different between a setting involving related dialects (as in most variationist studies of sound change) and a setting involving genetically and typologically distinct languages. Child vs adult language acquisition also matters, but this also does not diminish the role of social factors because part of the distinction between child and adult language acquisition is one that involves different social processes.

In recognizing that language matters, a major contribution of this study is in developing a model of how both macro-level and micro-level social factors are involved in initiating and in propagating sound change in a multilingual community in which linguistic distinctions are recognized. I have shown that just as the feature pool sets the starting point for possible linguistic changes, macro-level social forces set the starting point for the social context. Micro-level forces can either reinforce or challenge any tendencies set by these macro-level forces. To illustrate, I discussed the sociolinguistic history of Toronto Cantonese speakers from both a macro-level and a micro-level perspective in Chapter 4. The macro-level social forces involved account for why Hong Kong has become the center of Cantonese. For instance, Hong Kong's growth throughout the 20th century as a British colony led to the recognition of the Hong Kong variety as the prestige variety of Cantonese. Macro-level social forces such as changes in Canadian immigration laws and the 1997 handover to China are all important in accounting for why there was migration of Cantonese speakers from Hong Kong to Toronto. Also mentioned in Section 4.3 is the fact that many immigrants from Hong Kong were relatively affluent and opened Chinese businesses in the Greater Toronto Area. This was important in creating social conditions that have facilitated inter-generational maintenance of Cantonese in Toronto. This social context is also what makes contact-induced change from Toronto English to Cantonese possible. The keyword here is "possible". The macro-sociological context is not a deterministic force in sound change. As shown in the results from this study, only some GEN 2 speakers showed evidence for phonological change influenced by Toronto English. Many GEN 2 speakers did not change the vowel system by becoming fluent speakers of Toronto English. Instead, many GEN 2 speakers showed more overall conservatism in their vowel production patterns than younger Hong Kong speakers.

These results support a point made by Thomason that "it is not safe to assume that degree of cultural diffusion will correlate with degree of linguistic diffusion" (2001:196). As Thomason mentions, Montana Salish speakers have widely adopted European-American culture. Their language has clearly become endangered as a result but this has had little impact on the actual structure of Montana Salish. Thomason (2001) attributes the lack of contact-induced change to speaker attitudes about maintaining cross-linguistic differences. When Montana Salish speakers do introduce new words into their language, they prefer calquing over direct borrowing of loan words from English. A similar situation was observed in this dissertation. Rather than borrowing English loan words with Toronto English phonology intact, some GEN 2 speakers prefer to pronounce these words with Cantonese phonology. They use correspondence rules that they have developed from recognizing how many other English loan words are pronounced among GEN 1 speakers. They do not do this because they are unable to pronounce English loan words with Toronto English phonology. They do this because they believe that pronouncing these loanwords with Cantonese phonology is a part of speaking Cantonese.

What the Montana Salish and Toronto Cantonese examples illustrate is that a distinction between macro-level and micro-level social forces is important when describing the role of social factors in language change. In the Toronto Heritage Cantonese, we have an example of the overseas migration of a group of speakers to a different part of the world. This is arguably an extreme macro-level social change. Yet, most of the GEN 2 speakers interviewed showed lack of phonetic or phonological change. The macro-level social context only sets the possibilities and constraints while speakers can resist the possibilities created by the social context. The Montana Salish and Toronto Heritage Cantonese cases highlight the role of language attitudes and hence micro-level social factors. Macro-sociological forces are not deterministic because micro-level social factors can either challenge or support tendencies that arise from macro-sociological forces. It all depends on what individual speakers do, not on what societies per se do. Thus, both the Montana Salish and Toronto Heritage Cantonese examples illustrate cases in which contact-induced structural change is possible, but some speakers avoid cross-linguistic interference because of beliefs that the phonological structure and phonetic production patterns of Cantonese should be maintained. These beliefs that there exists a correct way of speaking Cantonese appear to be widespread even among speakers who show the most structural influence from English.

The relationship between macro-sociological forces and micro-level use of language in social interaction is one of the key differences between First Wave and Third Wave approaches to variation. While first wave studies have assumed that macro-level social forces affect micro-level language use, third wave studies have flipped around this relationship. Eckert says that "in the move from the first to the third wave of variation studies, the entire view of the relation between language and society has been reversed" (2012:97) Furthermore Eckert also says that contra first wave research, "patterns of variation do not simply unfold from the speaker's structural position in a system of production, but are part of the active—stylistic—production of social differentiation" (2012:98). Thus, macro-level social structures do not have a deterministic force on language use and linguistic structure as has already been discussed. Rather, speakers' use of linguistic variants in social interaction create meaning for the macro-sociological categories (ex: gender, social class) that have been the starting point for first wave studies. The findings from this dissertation, thus, show support for Third Wave ideas about the importance of micro-level social factors and about the problems of placing the explanatory burden on macro-level social forces by focusing on vowel variation in a heritage language. This is a type of contact setting that has not been previously addressed in Third Wave studies of vowel variation.

In supporting the overall argument that there is a difference between prototypical monolingual speakers and multilingual speakers who have acquired two or more distinct languages, I have also presented specific examples of how phonological structure can be accessible to the latter group. This also illustrates an example of micro-level social factors. The accessibility of phonology runs counter to Eckert & Labov who have recently claimed that "meaning accrues specifically to concrete sounds – to phonetic elements – and not to the phonological structures in which those sounds participate" (2017:467). I have shown that structure can, in fact, be accessible to speakers and that it can have social meaning. Speaking Cantonese, for example, means using Cantonese phonology even for pre-existing and new English loan words. The productivity of correspondence rules that I illustrated in Chapter 7 shows how Cantonese phonological maintenance is possible even if it means introducing more English loan words.

Eckert & Labov (2017) have also said that "Mergers, near mergers, splits, chain shifts and parallel shifts are not generally objects of social perception, conscious or unconscious, and are motivated by more abstract principles of change" (2017:467) In Chapter 6, however, I presented an interview excerpt involving metalinguistic discussion of a (near)-merger. Thus, contra Eckert and Labov's (2017) claim, I have shown that such phonological changes can become objects of social perception. Interference is clearly possible, but at the same time stigmatized because it indicates lack of proficiency in Cantonese. The accessibility of phonological structure is possible because of early acquisition of two distinct languages. This makes metalinguistic awareness relatively high leading to multiple possible outcomes on the individual level. Thus, some speakers show lack of change while others show phonological change.

Highlighting how these constraints are violated is important because doing so reveals a fundamental flaw with the TD Model: It was designed based on communities of idealized monolingual speakers who are poor learners of second languages. Constraints are a problem in the bulk of sociolinguistic research because in the bulk of communities studied, speakers lack structurally distinct linguistic choices. Heritage speakers, however, lack these constraints. This makes possible both interference and lack of interference. This is a point that is important in both the TK and VC Models. These models of contact-induced change are, thus, a more helpful starting point for pushing variationist research on sound change beyond its monolingually oriented core.

In addition to contributing to debates about 'language' vs. 'external forces', this dissertation also contributes to debates about the individual vs. the social in sound change. This is related to another major difference between first and third wave approaches. As Eckert says, "The emphasis on stylistic practice in the third wave places speakers not as passive and stable carriers of dialect, but as stylistic agents, tailoring linguistic styles in ongoing and lifelong projects of self-construction and differentiation" (2012:97–98). Here, Eckert is highlighting the importance of what individual speakers do rather than what societies or communities do. This emphasis on the individual, however, is similar to Thomason's (2007) discussion of deliberate change. The focus on the individual also works in tandem with a shift in focus from macro-level to micro-level social phenomena. This is a point I discuss further in the next section.

8.2 HERITAGE LANGUAGES OUT IN THE WILD

This dissertation also contributes to the relatively new field of heritage language phonetics and phonology. Much of this research has developed within a psycholinguistic rather than a sociolinguistic framework. The contribution I make to this emerging research area is in illustrating how the use of spontaneous speech data (and hence data collected “out in the wild”) can lead to a different perspective of heritage language speech from what can be concluded based on psycholinguistic approaches. This makes it possible to observe some unique insights into heritage speaker phonetics and phonology that may not be as easily observable in controlled psycholinguistic studies. The results from this study also challenge claims made in earlier studies about the stability of the phonology of heritage speakers compared to monolingual speakers of the same language.

Nagy (2015) has suggested that different methodological approaches to the study of heritage language use can lead to different results. For instance, while attrition is widely reported in experimental studies, the comparative variationist approach adopted by Nagy (2015) shows lack of evidence for attrition for Pro-drop and VOT. One observation made in this dissertation that may not have been as easy to make using an experimental approach is that GEN 2 speakers sometimes use correspondence rules to integrate English vocabulary into Cantonese. This turned out to be essential to explaining why there was an overall lack of phonetic and phonological change. From this observation, I was able to explain how it has been possible for GEN 2 speakers to keep the vowel system largely intact even while there is evidence that Cantonese vocabulary size and Cantonese usage has declined. Through examples presented in Chapter 7, I showed how the use of correspondence rules has become a productive process that allows GEN 2 speakers to introduce new English vocabulary while maintaining Cantonese phonology. Although this dissertation was not intended to be a code-switching study, it is clear from the use of correspondence rules and explicit metalinguistic discussions about Cantonese as a language loaded with English loan words pronounced with Cantonese phonology that further studies of code-switching and code-mixing behavior would be a fruitful area for future research.

The other major finding of this study is that not all GEN 2 speakers share the same Cantonese phonology. Polinsky and Kagan have said that “heritage speakers generally sound so native like – one could easily imagine that there would be no differences in phonological representations between the heritage language and the baseline, although that remains to be shown” (2007:378). The results from this study show that some GEN 2 speakers have innovated structural changes while others are relatively conservative. The structural changes include vowel splits and one (near)-merger. The use of spontaneous speech samples was helpful in identifying the specific vowels most susceptible to change and in identifying some of the factors that lead certain individual speakers into being more innovative than others. As discussed in Chapter 6, token counts of each individual word uttered in the interviews were tabulated making it possible to compile various descriptive statistics related to token frequencies of individual words and individual vowel categories across all three groups of speakers. Based on these descriptive statistics, it was possible to show a relationship between how frequent a vowel occurs in spontaneous speech and its susceptibility to structural change. For instance, two of the vowels that were shown to occur the least often in the corpus were also the two vowels involved in a merger. On the other end, two of the vowels (/ɔ/ and /a/) that occurred the most often are also two vowels that showed lack of change. Hong Kong data provided further support for explanations based on contact since the Hong Kong data shows lack of a /y/ vs. /u/ merger and a set of vowel shifts absent in Toronto.

In this dissertation, I have also been able to offer an explanation as to why some speakers show structural changes while many others do not. The factors that I showed that were the most predictive of the structural changes observed were factors that relate to speaker proficiency in a language. Gender, which is a common social variable included in many sociolinguistic analyses, failed to show statistical significance in accounting for anything. Similarly, age is also a social factor that accounts for change in many sociolinguistic studies, but the results from this study showed that it is important only in accounting for the apparent time fronting of /i/, which is not even a change unique to Toronto. Some EOQ factors predicted either /y/ retraction or /u/ fronting but none of these EOQ responses account for the splits in /ɛ/.

The two factors that consistently account for all of the structural changes observed were CAN % and ENG WC Scores. The importance of these factors underscores the importance of the individual. The discussion of the individual brings us back to where I left off in the previous section. Why do individual differences exist? Individual variation does not mean that social factors are irrelevant. We cannot ignore that becoming proficient in any language is itself a social process. Individual speakers need to be socialized into learning what Cantonese sounds like and in to learning the structural aspects of the language. The details involved are beyond the scope of the current project. What I have been able to show, however, is that research on heritage speaker phonetic and phonological variation is a topic that raises important issues in the interface between language and all the possible social forces that can have an effect on its use and structure. It also raises the issue of the individual vs. the social, which is another question that has been the subject of much debate. One point that is clear is that it would be a mistake to characterize heritage speakers as a homogenous group.

8.3 EXTENDING THE UNIFORMITARIAN PRINCIPLE

I end this dissertation by returning to the Uniformitarian Principle. That one could study the present with the goal of better understanding the past is what motivated Labov to study sound change in progress. Thanks to the pioneering work of Labov and his students, we now have a lot more ideas about how sound change could have developed in the past. I hope this dissertation has contributed to another piece of the puzzle by considering a modern and observable case of phonological development under intense contact.

Given the widespread loss of heritage languages after three generations in many communities across North America, one may be skeptical about the survival of Toronto Heritage Cantonese. This, however, should not diminish the importance of this dissertation as a contribution to research on sound change. As Simpson (2014) notes, even though language attrition has occurred in many communities across the world both past and present as a result of language shift to a dominant language, sometimes language shift stops. Not all of these languages become extinct. One example she mentions is the fate of English centuries after the Norman conquest. English was once a minority language just as Cantonese is today in the Toronto context, but it managed to survive and is now widely recognized as a global lingua franca.

Even if a language goes through a period of loss of speakers because of language shift to a dominant language, the socio-political status of a language can change even while changes introduced during periods of heavier L2 to L1 influence are transmitted to subsequent generations of speakers. History is full of other examples of intense contact situations that have led to various outcomes. This includes bilingual mixed languages, pidgins, and creoles. Heritage languages that have been maintained for more than three generations also exist such as Romani and Turkish dialects of Greek. In Toronto, five generations of heritage Ukrainian speakers have been recorded for the HLVC Project Corpus (Shkvorets 2015). Yet, the linguistic development of many of these languages has been controversial partly because of our inability to observe how they actually developed.

Without the ability to directly observe the past contact, it could be that we will never know for sure how many different languages developed. This is a weakness that Labov admits that we all have when he presents the historical paradox. He says, "the task of historical linguistics is to explain the differences between the past and the present; but to the extent that the past was different from the present, there is no way of knowing how different it was" (Labov 1994:21). The study of heritage languages in the present, however, may still be the closest observable analog available to researchers in the early 21st century to the development of languages such as the Turkish dialects of Greek, Romani, Media Lengua, Fon, and many others.

As Simpson says about such historic cases, "we do not yet have clear ways of deciding which changes are due to loss, borrowing, shift, restructuring and convergence under intense contact, and which are internal changes accelerated through intense contact" (2014:551). Thomason & Kaufman have said that "what is needed is research on current or recent contact situations that permit a more ambitious analysis of sociolinguistic context than we have attempted here" (1988:213). This dissertation has attempted to do exactly that. I hope that this dissertation has offered some ideas and that it inspires similar studies of other multilingual communities in the present so that we can better understand sound change and phonological development in a wider range of contact situations in the past.

APPENDIX A

LANGUAGE PERCENTAGE AND WORD COUNT SCORES

Table 70. GEN 2 Complete speaker summary

	CAN %	CAN Unique Word Count	ENG Unique Word Count
C2F24A	29.36%	356	1003
C2M21B	51.73%	308	615
C2F21C	67.72%	409	460
C2F22A	74.25%	486	523
C2M22A	79.50%	408	395
C2F41A	84.75%	465	233
C2F20A	85.30%	471	354
C2M21D	86.92%	432	281
C2M44A	89.77%	649	419
C2M27A	92.83%	404	137
C2M21C	94.00%	566	313
C2F21B	98.59%	715	100

APPENDIX B

MEAN F1 AND F2 VALUES FOR EACH INDIVIDUAL SPEAKER

Table 71. F1 and F2 Means, GEN 1, tense vowels

	/a/		/ɛ/		/ɔ/		/œ/		/i/		/y/		/u/	
Speakers	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
C1F50A	509	1357	428	1569	438	1161	430	1340	360	1780	360	1619	388	1194
C1F50B	508	1322	431	1529	424	1173	422	1336	373	1764	366	1626	377	1149
C1F54B	501	1377	427	1612	421	1176	430	1374	361	1831	356	1664	372	1208
C1F58A	515	1324	437	1552	451	1183	454	1324	375	1728	368	1587	385	1130
C1F78A	509	1341	427	1596	429	1138	429	1356	367	1713	365	1629	380	1136
C1F83A	503	1353	413	1621	442	1153	433	1367	361	1720	356	1586	368	1099
C1M46A	499	1352	430	1605	441	1162	435	1369	351	1759	340	1647	373	1183
C1M52A	506	1326	423	1601	448	1179	441	1355	345	1778	353	1647	380	1146
C1M52B	504	1339	419	1580	448	1194	439	1394	347	1763	352	1714	374	1124
C1M59A	490	1340	407	1664	439	1173	427	1412	324	1796	327	1754	348	1207
C1M61A	508	1329	434	1568	450	1139	443	1383	365	1694	367	1572	379	1147
C1M87A	504	1366	431	1611	431	1180	419	1315	376	1711	366	1653	393	1185

Table 72. F1 and F2 Means, GEN 2, tense vowels

	/a/		/ɛ/		/ɔ/		/œ/		/i/		/y/		/u/	
Speakers	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
C2F20A	499	1359	425	1572	445	1204	437	1390	341	1808	349	1620	369	1199
C2F21B	496	1324	447	1553	441	1166	459	1356	355	1751	349	1642	377	1141
C2F21C	501	1382	423	1596	446	1207	445	1332	343	1797	352	1627	350	1073
C2F22A	508	1348	431	1555	428	1176	422	1357	370	1767	362	1616	374	1170
C2F24A	500	1329	438	1610	427	1200	426	1365	358	1824	358	1514	354	1284
C2F41A	513	1280	434	1542	438	1187	434	1328	364	1773	361	1517	387	1175
C2M21B	497	1333	419	1632	430	1171	418	1400	354	1764	332	1565	369	1291
C2M21C	505	1332	431	1581	426	1181	428	1344	353	1792	358	1682	397	1254
C2M21D	512	1326	445	1560	441	1161	445	1363	369	1698	373	1573	372	1154
C2M22A	505	1410	417	1603	411	1197	410	1448	363	1915	366	1548	375	1199
C2M27A	510	1262	428	1583	436	1174	421	1426	363	1721	362	1645	383	1109
C2M44A	510	1316	439	1598	444	1118	438	1383	362	1689	363	1607	377	1125

Table 73. F1 and F2 Means, HK, tense vowels

	/a/		/ɛ/		/ɔ/		/œ/		/i/		/y/		/u/	
Speakers	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
CXF16A	497	1364	422	1638	441	1192	435	1432	345	1807	335	1621	366	1117
CXF19A	512	1317	421	1641	421	1180	424	1390	364	1845	356	1567	354	1181
CXF43A	510	1379	424	1635	424	1182	431	1376	352	1812	339	1644	355	1147
CXF49A	514	1332	418	1647	421	1168	418	1348	359	1804	354	1557	359	1194
CXF77A	508	1369	417	1632	414	1157	427	1387	361	1791	356	1691	352	1073
CXM20A	511	1363	422	1652	427	1203	428	1372	344	1851	341	1694	349	1136
CXM27A	504	1429	432	1696	433	1186	454	1450	336	1878	325	1678	353	1107
CXM52A	499	1333	421	1652	424	1164	426	1352	349	1815	347	1680	365	1202

Table 74. F1 and F2 Means, GEN 1, lax vowels

	/ɐ/		/ɵ/		/ɪ/		/ʊ/	
Speakers	F1	F2	F1	F2	F1	F2	F1	F2
C1F50A	444	1332	397	1372	360	1662	407	1215
C1F50B	444	1328	406	1390	373	1626	394	1149
C1F54B	445	1382	379	1386	361	1761	383	1207
C1F58A	453	1298	414	1346	375	1626	409	1206
C1F78A	461	1353	396	1481	367	1654	396	1180
C1F83A	463	1356	408	1379	361	1692	405	1139
C1M46A	450	1334	411	1356	351	1658	411	1195
C1M52A	458	1354	422	1348	345	1622	409	1198
C1M52B	461	1329	418	1374	347	1642	424	1201
C1M59A	452	1372	412	1354	324	1705	424	1185
C1M61A	460	1356	402	1382	365	1645	408	1141
C1M87A	452	1364	427	1472	376	1698	416	1204

Table 75. F1 and F2 Means, GEN 2, lax vowels

	/ɐ/		/ɵ/		/ɪ/		/ʊ/	
Speakers	F1	F2	F1	F2	F1	F2	F1	F2
C2F20A	447	1366	396	1444	382	1726	408	1209
C2F21B	464	1334	414	1376	403	1636	416	1183
C2F21C	449	1377	424	1393	392	1654	405	1173
C2F22A	463	1345	384	1405	398	1636	411	1223
C2F24A	471	1339	411	1400	395	1627	414	1282
C2F41A	459	1316	390	1427	393	1680	405	1205
C2M21B	462	1329	407	1526	403	1672	414	1226
C2M21C	444	1324	398	1359	383	1708	406	1225
C2M21D	461	1321	415	1363	397	1620	406	1162
C2M22A	440	1379	392	1395	390	1724	400	1182
C2M27A	462	1314	408	1408	398	1605	413	1210
C2M44A	464	1341	399	1447	389	1635	419	1151

Table 76. F1 and F2 Means, HK, lax vowels

	/ɐ/		/ɵ/		/ɪ/		/ʊ/	
Speakers	F1	F2	F1	F2	F1	F2	F1	F2
CXF16A	449	1369	391	1530	418	1654	397	1278
CXF19A	454	1347	394	1439	394	1637	401	1226
CXF43A	461	1361	408	1308	397	1690	390	1194
CXF49A	465	1334	391	1302	391	1662	390	1200
CXF77A	453	1396	389	1439	391	1653	391	1200
CXM20A	466	1380	422	1383	410	1642	404	1208
CXM27A	460	1390	431	1406	411	1636	399	1201
CXM52A	453	1376	393	1377	399	1631	397	1210

APPENDIX C

PILLAI SCORES FOR EACH INDIVIDUAL SPEAKER

Table 77. Pillai Scores for /y/ vs. /u/

GEN 1 Speakers	Pillai Score	**GEN 2 Speakers**	Pillai Score	**HK Speakers**	Pillai Score
C1F50A	0.869	**C2F20A**	0.869	**CXF16A**	0.863
C1F50B	0.834	**C2F21B**	0.926	**CXF19A**	0.887
C1F54B	0.8	**C2F21C**	0.901	**CXF43A**	0.935
C1F58A	0.88	**C2F22A**	0.852	**CXF49A**	0.941
C1F78A	0.913	**C2F24A**	0.705	**CXF77A**	0.944
C1F83A	0.895	**C2F41A**	0.747	**CXM20A**	0.924
C1M46A	0.798	**C2M21B**	0.854	**CXM27A**	0.944
C1M52A	0.898	**C2M21C**	0.763	**CXM52A**	0.926
C1M52B	0.935	**C2M21D**	0.875		
C1M59A	0.927	**C2M22A**	0.565		
C1M61A	0.832	**C2M27A**	0.926		
C1M87A	0.869	**C2M44A**	0.897		

Table 78. Pillai Scores for [i] vs. [ɪ]

GEN 1 Speakers	Pillai Score	**GEN 2 Speakers**	Pillai Score	**HK Speakers**	Pillai Score
C1F50A	0.529	**C2F20A**	0.494	**CXF16A**	0.682
C1F50B	0.487	**C2F21B**	0.525	**CXF19A**	0.664
C1F54B	0.206	**C2F21C**	0.577	**CXF43A**	0.613
C1F58A	0.533	**C2F22A**	0.551	**CXF49A**	0.614
C1F78A	0.292	**C2F24A**	0.608	**CXF77A**	0.449
C1F83A	0.274	**C2F41A**	0.46	**CXM20A**	0.778
C1M46A	0.5	**C2M21B**	0.546	**CXM27A**	0.814
C1M52A	0.676	**C2M21C**	0.352	**CXM52A**	0.596
C1M52B	0.72	**C2M21D**	0.512		
C1M59A	0.623	**C2M22A**	0.454		
C1M61A	0.398	**C2M27A**	0.413		
C1M87A	0.148	**C2M44A**	0.476		

Table 79. Pillai Scores for /ɛ/ in open syllable vs. nasal coda context

GEN 1 Speakers	Pillai Score	**GEN 2 Speakers**	Pillai Score	**HK Speakers**	Pillai Score
C1F50A	0.001	**C2F20A**	0.393	**CXF16A**	0.104
C1F50B	0.068	**C2F21B**	0.06	**CXF19A**	0.176
C1F54B	0.243	**C2F21C**	0.164	**CXF43A**	0.234
C1F58A	0.279	**C2F22A**	0.081	**CXF49A**	0.427
C1F78A	0.039	**C2F24A**	0.535	**CXF77A**	0.138
C1F83A	0.311	**C2F41A**	0.452	**CXM20A**	0.448
C1M46A	0.169	**C2M21B**	0.467	**CXM27A**	0.262
C1M52A	0.162	**C2M21C**	0.209	**CXM52A**	0.29
C1M52B	0.217	**C2M21D**	0.057		
C1M59A	0.169	**C2M22A**	0.517		
C1M61A	0.029	**C2M27A**	0.277		
C1M87A	0.002	**C2M44A**	0.121		

Table 80. Pillai Scores for /ɔ/ in open syllable vs. closed syllable context

GEN 1 Speakers	Pillai Score	**GEN 2 Speakers**	Pillai Score	**HK Speakers**	Pillai Score
C1F50A	0.114	**C2F20A**	0.2	**CXF16A**	0.133
C1F50B	0.195	**C2F21B**	0.169	**CXF19A**	0.214
C1F54B	0.124	**C2F21C**	0.077	**CXF43A**	0.178
C1F58A	0.321	**C2F22A**	0.273	**CXF49A**	0.192
C1F78A	0.144	**C2F24A**	0.394	**CXF77A**	0.316
C1F83A	0.25	**C2F41A**	0.28	**CXM20A**	0.249
C1M46A	0.157	**C2M21B**	0.138	**CXM27A**	0.103
C1M52A	0.307	**C2M21C**	0.242	**CXM52A**	0.358
C1M52B	0.256	**C2M21D**	0.192		
C1M59A	0.283	**C2M22A**	0.22		
C1M61A	0.161	**C2M27A**	0.081		
C1M87A	0.124	**C2M44A**	0.078		

BIBLIOGRAPHY

Adank, Patti.; Roel Smits.; and Roeland Van Hout. 2004. A comparison of vowel normalization procedures for language variation research. The Journal of the Acoustical Society of America 116.3099–3107.

Bacon-Shone, John.; and Kingsley Bolton. 1998. Charting Multilingualism: Language Censuses and Language Surveys in Hong Kong. Language in Hong Kong at century's end, ed. by Martha Carswell Pennington, 43–90. Hong Kong: Hong Kong University Press.

Baker, Wendy.; and Pavel Trofimovich. 2005. Interaction of Native- and Second-Language Vowel System(s) in Early and Late Bilinguals. Language and Speech 48.1–27.

Barrie, Mike. 2003. Contrast in Cantonese vowels. Toronto Working Papers in Linguistics 20.1–19.

Bauer, Robert. 1982. Cantonese Sociolinguistic Patterns: Correlating Social Characteristics of Speakers with Phonological Variables in Hong Kong Cantonese. University of California, Berkeley Ph.d. dissertation. http://escholarship.org/uc/item/1sn2f75h.

Bauer, Robert S. 1985. The expanding syllabary of Hong Kong cantonese. Cahiers de linguistique - Asie orientale 14.99–111. doi:10.3406/clao.1985.1169.

Bauer, Robert S.; and Paul K. Benedict. 1997. Modern Cantonese Phonology. Trends in Linguistics 102. Berlin: Mouton de Gruyter.

Berridge, Joe. 1995. Is Toronto different from U.S. cities? Was it planned that way? On the occasion of a major planning conference in Toronto, urban planner Joe Berridge tries to explain the city to U.S. delegates. "Torontonians are the only people on Earth who believe that sin is a failure of education, and that sex is a branch of public health." The Globe and Mail (Canada) April 10, 1995.

Bloomfield, Leonard. 1933. Language. New York: H. Holt and Company.

Boberg, Charles. 2008. Regional Phonetic Differentiation in Standard Canadian English. Journal of English Linguistics 36.129–154. doi:10.1177/0075424208316648.

Bucholtz, Mary. 2003. Sociolinguistic nostalgia and the authentication of identity. Journal of Sociolinguistics 7.398–416. doi:10.1111/1467-9481.00232.

Bucholtz, Mary. 2009. Styles and stereotypes: Laotian American girls' linguistic negotiation of identity. Beyond yellow English: toward a linguistic anthropology of Asian Pacific America, ed. by Angela Reyes and Adrienne Lo, 21–42. Oxford studies in sociolinguistics. Oxford ; New York: Oxford University Press.

Calvet, Louis-Jean. 2003. Reflections on the Origins of Sociolinguistics in Europe. Sociolinguistics: the essential readings, ed. by Christina Bratt Paulston and G. Richard Tucker. Malden, MA: Blackwell Pub.

Campbell, Lyle. 1999. Historical linguistics: an introduction. 1st MIT Press ed. Cambridge, Mass: MIT Press.

Carroll, John M. 2007. A concise history of Hong Kong. Critical issues in history. Lanham: Rowman & Littlefield.

Chan, Arlene. 2011. The Chinese in Toronto from 1878: from outside to inside the circle. Toronto: Dundurn Press.

Chang, Charles B.; Yao Yao.; Erin F. Haynes.; and Russell Rhodes. 2011. Production of phonetic and phonological contrast by heritage speakers of Mandarin. The Journal of the Acoustical Society of America 129.3964–3980. doi:10.1121/1.3569736.

Chen, Katherine.; and Gray Carper. 2005. Multilingual Hong Kong: Duo yu Xianggang. Episode I, Present yi ge project[Videorecording] =. Films for the Humanities & Sciences.

Chen, Katherine Hoi Ying. 2005. The social distinctiveness of two code-mixing styles in Hong Kong. Proceedings of the 4th international symposium on bilingualism, 527–541. http://www.lingref.com/isb/4/039ISB4.PDF.

Chen, Katherine Hoi Ying. 2008. Positioning and repositioning: Linguistic practices and identity negotiation of overseas returning bilinguals in Hong Kong. Multilingua - Journal of Cross-Cultural and Interlanguage Communication 27.57–75. doi:10.1515/MULTI.2008.004.

Cheshire, Jenny.; Paul Kerswill.; Sue Fox.; and Eivind Torgersen. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. Journal of Sociolinguistics 15.151–196. doi:10.1111/j.1467-9841.2011.00478.x.

Cheung, Yat-shing. 1985. Power, Solidarity, and Luxury in Hong Kong: A Sociolinguistic Study. Anthropological Linguistics 27.190–203.

Clarke, Sandra.; Ford Elms.; and Amani Youssef. 1995. The third dialect of English: Some Canadian evidence. Language Variation and Change 7.209–228. doi:10.1017/S0954394500000995.

Coupland, Nikolas. 2014. Language Change, Social Change, Sociolinguistic Change: A Meta-Commentary. Journal of Sociolinguistics 18.277–286. doi:10.1111/josl.12077.

Coupland, Nikolas. 2016. Labov, vernacularity and sociolinguistic change. Journal of Sociolinguistics 20.409–430. doi:10.1111/josl.12191.

Cui, Naomi. 2014. Labber. Toronto.

Cui, Naomi.; Minyi Zhu.; Vina Law.; Holman Tse.; and Naomi Nagy. 2014. Exploring automated formant analysis for comparative variationist study of Heritage Cantonese and English. Presentation at Change and Variation in Canada / Changement et Variation au Canada (CVC 8). Queen's University, Kingston, ON, Canada. http://d-scholarship.pitt.edu/26261/.

DeFrancis, John. 1984. The Chinese language: fact and fantasy. Honolulu: University of Hawaii Press.

Dong, Lorraine.; and Marlon K. Hom. 1980. Chinatown Chinese: The San Francisco Dialect. Amerasia Journal 7.1–29.

Dubois, Sylvie.; and Barbara Horvath. 1999. When the music changes, you change too: Gender and language change in Cajun English. Language Variation and Change 11.287–313. doi:10.1017/S0954394599113036.

Eckert, Penelope. 2003. Elephants in the room. Journal of Sociolinguistics 7.392–397.

Eckert, Penelope. 2008. Where do ethnolects stop? International Journal of Bilingualism 12.25–42. doi:10.1177/13670069080120010301.

Eckert, Penelope. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. Annual Review of Anthropology 41.87–100. doi:10.1146/annurev-anthro-092611-145828.

Eckert, Penelope.; and William Labov. 2017. Phonetics, Phonology and Social Meaning. Journal of Sociolinguistics 21.467–496. doi:10.1111/josl.12244.

Epps, Patience.; John Huehnergard.; and Na'ama Pat-El. 2013. Introduction: Contact Among Genetically Related Languages. Journal of Language Contact 6.209–219. doi:10.1163/19552629-00602001.

Evans, Stephen. 2014. The decline and fall of English in Hong Kong's Legislative Council. Journal of Multilingual and Multicultural Development 35.479–496. doi:10.1080/01434632.2013.873802.

Flege, James E. 1995. Second-language speech learning: Theory, findings, and problems. Speech perception and linguistic experience, ed. by Winifred Strange, 233–277.

Fought, Carmen. 2003. Chicano English in context. Palgrave Macmillan.

Garrett, Andrew. 2014. Sound change. The Routledge Handbook of Historical Linguistics, ed. by Claire Bowern and Bethwyn Evans, 227–248. London, UNITED KINGDOM: Routledge.

Godson, Linda. 2004. Vowel Production in the Speech of Western Armenian Heritage Speakers. Heritage Language Journal 2.1–26.

Gooden, Shelome (in press) Festschrift in honor of John R. Rickford. In the Fisherman's Net. Language Contact in a sociolinguistics context. In Renee Blake and Isabelle Buchstaller (eds). *The Routledge Companion to the Work of John R. Rickford*. Routledge

Gorman, Kyle.; Jonathan Howell.; and Michael Wagner. 2011. Prosodylab-aligner: A tool for forced alignment of laboratory speech. Canadian Acoustics 39.192–193.

Guy, Gregory R. 1990. The Sociolinguistic Types of Language Change. Diachronica 7.47–67. doi:10.1075/dia.7.1.04guy.

Hall-Lew, Lauren. 2009. Ethnicity and phonetic variation in a San Francisco neighborhood. Stanford, CA: Stanford University unpublished phd dissertation.

Hall-Lew, Lauren. 2010. Improved representation of variance in measures of vowel merger. Proceedings of Meetings on Acoustics 159ASA, 9:060002. ASA.

Harrison, Godfrey.; and Lydia So. 1997. The Background to Language Change in Hong Kong. One country, two systems, three languages: a survey of changing language use in Hong Kong, ed. by Sue Wright and Helen Kelly-Holmes, 8–17. Clevedon, England: Multilingual Matters.

Hay, Jennifer.; Paul Warren.; and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. Journal of Phonetics 34.458–484.

Herold, Ruth. 1990. Mechanisms of merger: The implementation and distribution of the low back merger in eastern Pennsylvania. United States -- Pennsylvania: University of Pennsylvania ph.d. dissertation.

Herold, Ruth. 1997. Solving the actuation problem: Merger and immigration in eastern Pennsylvania. Language variation and change 9.165–189.

Hockett, Charles Francis. 1958. A course in modern linguistics. New York: The Macmillan Company.

Hoffman, Michol F. 2010. The Role of Social Factors in the Canadian Vowel Shift: Evidence from Toronto. American Speech 85.121–140. doi:10.1215/00031283-2010-007.

Hoffman, Michol F.; and James A. Walker. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. Language Variation and Change 22.37–67. doi:10.1017/S0954394509990238.

Johanson, Lars. 2008. Remodeling grammar: Copying, conventionalization, grammaticalization. Language contact and contact languages, ed. by Peter Siemund and Noemi Kintana, 7:61–79. Hamburg studies on multilingualism. Amsterdam: John Benjamins Pub. Co.

Johnson, Daniel Ezra. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. Language and Linguistics Compass 3.359–383.

Johnstone, Barbara. 2013. Speaking Pittsburghese: The story of a dialect. Oxford University Press.

Joseph, John E. 1997. English in Hong Kong: Emergence and Decline. One country, two systems, three languages: a survey of changing language use in Hong Kong, ed. by Sue Wright and Helen Kelly-Holmes, 60–73. Clevedon, England ; Philadelphia: Multilingual Matters.

Kiesling, Scott F. 2005. Variation, stance and style: Word-final-er, high rising tone, and ethnicity in Australian English. English World-Wide 26.1–42. doi:10.1075/eww.26.1.02kie.

Kiesling, Scott F.; and Marc Wisnosky. 2003. Competing norms, heritage prestige, and/aw/-monophthongization in Pittsburgh. poster presented at New Ways of Analyzing Variation (NWAV) 32.

Kiparsky, Paul. 2016. Labov, Sound Change, and Phonological Theory. Journal of Sociolinguistics 20.464–488. doi:10.1111/josl.12196.

Labov, William. 1972. Sociolinguistic Patterns. University of Pennsylvania Press.

Labov, William. 1984. Field methods of the project on linguistic change and variation. Language in use: readings in sociolinguistics, ed. by John Baugh and Joel Sherzer, 28–53. Englewood Cliffs, NJ: Prentice Hall.

Labov, William. 1994. Principles of linguistic change, Volume 1. Oxford, UK; Cambridge, MA: Blackwell.

Labov, William. 2001. Principles of linguistic change. Vol. 2, Vol. 2,. Oxford: Blackwell.

Labov, William. 2007. Transmission and Diffusion. Language 83.344–387.

Labov, William. 2008. Mysteries of the substrate. Social Lives in Language–Sociolinguistics and multilingual speech communities: Celebrating the work of Gillian Sankoff., 315–326. Amsterdam, The Netherlands/Philadelphia, Pennsylvania: John Benjamins Publishing Company.

Labov, William. 2011. Principles of linguistic change, cognitive and cultural factors. Vol. 3. John Wiley & Sons.

Labov, William.; Sharon Ash.; and Charles Boberg. 2006. The Atlas of North American English: Phonetics, Phonology and Sound Change. Walter de Gruyter.

Labov, William.; Malcah Yaeger.; and Richard Steiner. 1972. A quantitative study of sound change in progress. Vol. 1. US Regional Survey.

Law, Danny. 2013. Inherited similarity and contact-induced change in Mayan Languages. Journal of Language Contact 6.271–299. doi:10.1163/19552629-00602004.

Lee, Jackson L. 2014. The Representation of Contour Tones in Cantonese. Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society 38.272–287.

Lee, Thomas. 1983. The vowel system in two varieties of Cantonese. UCLA Working Papers in Phonetics 57.97–114.

Leung, Cheung-Shing Samuel.; and Yuen-Fan Lornita Wong. 1997. The (Un)Changing Role of Mandarin Chinese in Language Education in Hong Kong. One country, two systems, three languages: a survey of changing language use in Hong Kong, ed. by Sue Wright and Helen Kelly-Holmes, 34–45. Clevedon, England: Multilingual Matters.

Leung, Genevieve Yuek-Ling. 2012. Hoisan-wa reclaimed: Chinese American language maintenance and language ideology in historical and contemporary sociolinguistic perspective. Philadelphia: University of Pennsylvania. http://repository.upenn.edu/dissertations/AAI3542824/.

Lynch, Andrew.; and Maria Polinsky. 2018. Native speakerhood in heritage language research. Connecting across languages and cultures: a heritage language festschrift in honor of Olga Kagan, ed. by Olga Kagan, Susan Bauckus, and Susan Kresin, 143–162. Bloomington, Indiana: Slavica Publishers.

Lyskawa, Paulina.; Ruth Maddeaux.; Emilia Melara.; and Naomi Nagy. 2016. Heritage Speakers Follow All the Rules: Language Contact and Convergence in Polish Devoicing. Heritage Language Journal 13.219.

Mack, Molly. 1990. Phonetic transfer in a French-English bilingual child. Language attitudes and language conflict.107–124.

Maddieson, Ian. 2013. Vowel Quality Inventories. The World Atlas of Language Structures Online, ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/2.

Man, Evelyn Yee-fun. 2006. First language use and language behavior of Chinese students in Toronto, Canada. Heritage Language Development : Focus on East Asian Immigrants, ed.

by Kimi Kondo-Brown, 209–241. Amsterdam, NLD: John Benjamins Publishing Company.

Matthews, Stephen.; and Virginia Yip. 2011. Cantonese: A Comprehensive Grammar. Routledge.

Mithun, Marianne. 2013. Challenges and Benefits of Contact among Relatives: Morphological Copying. Journal of Language Contact 6.243–270. doi:10.1163/19552629-00602003.

Montrul, Silvina. 2015. The acquisition of heritage languages. Cambridge University Press.

Mufwene, Salikoko S. 2001. The ecology of language evolution. Cambridge University Press.

Nagy, Naomi. 2011. A Multilingual Corpus to Explore Variation in Language Contact Situations. Rassegna Italiana di Linguistica Applicata 43.65–84.

Nagy, Naomi. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. Lingua. doi:10.1016/j.lingua.2014.04.012.

Nagy, Naomi. 2016. Heritage languages as new dialects. The future of dialects: Selected papers from Methods in Dialectology XV , ed. by Marie-Hélène Côté, Remco Knooihuizen, and John Nerbonne, 15–34. Language Variation. Berlin: Language Science Press. http://langsci-press.org/catalog/book/81.

Nagy, Naomi.; Joanna Chociej.; and Michol F. Hoffman. 2014. Analyzing Ethnic Orientation in the quantitative sociolinguistic paradigm. Language & Communication 35. New perspectives on linguistic variation and ethnic identity in North America.9–26. doi:10.1016/j.langcom.2013.11.002.

Nagy, Naomi.; Yoonjung Kang.; Alexei Kochetov.; and James Walker. 2009. Heritage Languages in Toronto: A New Project. Paper presented at the Heritage Language Workshop, University of Toronto.

Nagy, Naomi.; and Miriam Meyerhoff. 2008. Introduction: Social Lives in Language. Social lives in language--sociolinguistics and multilingual speech communities: celebrating the work of Gillian Sankoff, 1–17. Amsterdam; Philadelphia: John Benjamins.

Nurse, Derek.; and Thomas J. Hinnebusch. 1993. Swahili and Sabaki: a linguistic history. University of California publications in linguistics v. 121. Berkeley: University of California Press.

Nycz, Jennifer.; and Lauren Hall-Lew. 2015. Best practices in measuring vowel merger. Proceedings of Meetings on Acoustics 20.060008. doi:10.1121/1.4894063.

Pannu, Jasbir. 1998. Language Choice and Identity: the World of the Hong Kong Indian Adolescent. Language in Hong Kong at century's end, ed. by Martha Carswell Pennington, 219–242. Hong Kong: Hong Kong University Press.

Peters, Andrew.; and Holman Tse. 2016. Evaluating the efficacy of Prosody-lab Aligner for a study of vowel variation in Cantonese. Columbus, United States of America: University of Pittsburgh. http://u.osu.edu/wicl/files/2016/03/wicl3_abstract_peters_tse-1vhqui3.pdf.

Pierson, Herbert D. 1998. Societal Accommodation to English and Putonghua in Cantonese-Speaking Hong Kong. Language in Hong Kong at century's end, ed. by Martha Carswell Pennington, 91–111. Hong Kong: Hong Kong University Press.

Polinsky, Maria. 2011. Heritage Languages. Linguistics. doi:10.1093/obo/9780199772810-0067.

Polinsky, Maria.; and Olga Kagan. 2007. Heritage Languages: In the 'Wild' and in the Classroom. Language and Linguistics Compass 1.368–395. doi:10.1111/j.1749-818X.2007.00022.x.

Romaine, Suzanne. 2012. Types of bilingual and multilingual speech communities. The Handbook of Bilingualism and Multilingualism, 2nd Edition, ed. by Tej K. Bhatia and William C. Ritchie. 2nd ed. Wiley-Blackwell.

Ronquest, Rebecca E. 2013. An acoustic examination of unstressed vowel reduction in Heritage Spanish. Selected proceedings of the 15th hispanic linguistics symposium, ed. by Chad Howe, Sarah E. Blackwell, and Margaret Lubbers Quesada, 157–171. http://www.lingref.com/cpp/hls/15/paper2882.pdf.

Rosenfelder, Ingrid.; Joe Fruehwald.; Keelan Evanini.; and Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite. http://fave.ling.upenn.edu.

Saadah, Eman. 2011. The production of Arabic vowels by English L2 learners and heritage speakers of Arabic. University of Illinois at Urbana-Champaign ph.d. dissertation. https://www.ideals.illinois.edu/handle/2142/24104.

Sankoff, Gillian. 2013. Linguistic Outcomes of Bilingualism. The handbook of language variation and change, ed. by J. K Chambers and Natalie Schilling-Estes, 501–518.

Scovel, Thomas. 2000. A CRITICAL REVIEW OF THE CRITICAL PERIOD RESEARCH. Annual Review of Applied Linguistics 20.213–223. doi:10.1017/S0267190500200135.

Setter, Jane.; Cathy S. P. Wong.; and Brian Hok-Shing Chan. 2010. Hong Kong English. Dialects of English. Edinburgh: Edinburgh University Press.

Shkvorets, Maksym. 2015. The Loss of Reflexive Possessive Pronouns in Heritage Ukrainian. University of Ottawa.

Shuy, Roger W. 2003. A Brief History of American Sociolinguistics 1949-1989. Sociolinguistics: the essential readings, ed. by Christina Bratt Paulston and G. Richard Tucker. Malden, MA: Blackwell Pub.

Siemiatycki, Myer.; Tim Rees.; Roxana Ng.; and Khan Rahi. 2003. Integrating Community Diversity in Toronto. The World in a City, ed. by Paul Anisef and C. Michael Lanphier. University of Toronto Press.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. Phonology 9.289–328.

Simpson, Jane. 2014. Language attrition and language change. The Routledge Handbook of Historical Linguistics, ed. by Claire Bowern and Bethwyn Evans, 537–554. London, UNITED KINGDOM: Routledge.

Sloetjes, Han.; and Peter Wittenburg. 2008. Annotation by category – ELAN and ISO DCR. Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008), 816–820.

Smits, Caroline. 1998. Two Models for the Study of Language Contact: A Psycho-Linguistic Perspective versus a Socio-Cultural Perspective. Historical Linguistics 1997, 377–390.

Snow, Donald B. 2004. Cantonese as written language: the growth of a written Chinese vernacular. Hong Kong: Hong Kong University Press.

Stewart, Jesse. 2014. A Comparative Analysis of Media Lengua and Quichua Vowel Production. Phonetica 71.159–182. doi:10.1159/000369629.

Stewart, Jesse. 2018. Vowel perception by native Media Lengua, Quichua, and Spanish speakers. Journal of Phonetics 71.177–193. doi:10.1016/j.wocn.2018.08.005.

Sturtevant, Edgar H. 1947. An Introduction To Linguistic Science. Yale University Press. http://archive.org/details/introductiontoli032816mbp.

Sung, Chit Cheung Matthew. 2015. Hong Kong English: Linguistic and Sociolinguistic Perspectives. Language and Linguistics Compass 9.256–270. doi:10.1111/lnc3.12142.

Tagliamonte, Sali A. 2011. Variationist Sociolinguistics: Change, Observation, Interpretation. Hoboken, GB: Wiley-Blackwell.

Taylor, Bill. 2007. The Language Quilt. Toronto Star (Canada).

Thomas, Erik.; and Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. http://lingtools.uoregon.edu/norm/norm1.php.

Thomas, Erik R. 2011. Sociophonetics: an introduction. Basingstoke, Hampshire; New York: Palgrave Macmillan.

Thomason, Sarah G. 2007. Language Contact and Deliberate Change. Journal of Language Contact 1.41–62. doi:10.1163/000000007792548387.

Thomason, Sarah Grey. 2001. Language Contact. Georgetown University Press.

Thomason, Sarah Grey.; and Terrence Kaufman. 1988. Language contact, creolization, and genetic linguistics. University of California Press.

Thompson, Richard H. 1989. Toronto's Chinatown: the changing social organization of an ethnic community. Vol. 29. Immigrant communities & ethnic minorities in the United States & Canada. New York, NY: AMS Press.

Tsang, Chui-Lim. 1984. Sociolinguistic Considerations in Selecting the Language of Instruction in a Bilingual Programme. New papers on Chinese language use, ed. by Beverly Hong, 45–58. Canberra: Contemporary China Centre, Research School of Pacific Studies, Australian National University.

Tsang, Steve. 2007. Modern History of Hong Kong, A. London, GB: I.B.Tauris.

Tse, Holman. 2016a. Contact-induced splits in Toronto Heritage Cantonese mid-vowels. Linguistica Atlantica 35.133–155.

Tse, Holman. 2016b. Variation and change in Toronto heritage Cantonese. Asia-Pacific Language Variation 2.124–156. doi:10.1075/aplv.2.2.02tse.

van Coetsem, Frans. 1988. Loan phonology and the two transfer types in language contact. Dordrecht, Holland; Providence R.I., U.S.A.: Foris Publications.

van Coetsem, Frans. 1990. Thomason and Kaufman: Language contact, creolization, and genetic linguistics (Book Review). Language in Society 19.260–268.

van Coetsem, Frans. 2000. A general and unified theory of the transmission process in language contact. Heidelberg: Winter.

Walker, James.; Naomi Nagy.; Michol Hoffman.; and Holman Tse. 2018. What Are Ethnolects Made of? Auckland, New Zealand.

Wang, W S Y. 1979. Language Change A Lexical Perspective. Annual Review of Anthropology 8.353–371. doi:10.1146/annurev.an.08.100179.002033.

Wang, William S.-Y. 1969. Competing Changes as a Cause of Residue. Language 45.9–25.

Weinreich, Uriel.; William Labov.; and Marvin Herzog. 1968. Empirical foundations for a theory of language change. Directions for historical linguistics: A symposium, ed. by Winfred Philipp Lehmann and Yakov Malkiel, 95–188. University of Texas Press.

Winford, Donald. 2003. An introduction to contact linguistics. Malden, Mass.: Blackwell Pub.

Winford, Donald. 2007. Some Issues in the Study of Language Contact. Journal of Language Contact 1.22–40. doi:10.1163/000000007792548288.

Wong, Amy Wing-mei.; and Lauren Hall-Lew. 2014. Regional Variability and Ethnic Identity: Chinese Americans in New York City and San Francisco. Language & Communication 35.27.

Wong, Wai Yi Peggy. 2006. Syllable fusion in Hong Kong Cantonese connected speech. The Ohio State University ph.d. dissertation.

Yang, Jing. 2014. Acoustic properties of vowel production in Mandarin-English bilingual and corresponding monolingual children. The Ohio State University.

Yip, Moira. 1996. Lexicon Optimization in Languages without Alternations. Rutgers Optimality Archive. doi:10.7282/T33X84P0. https://rucore.libraries.rutgers.edu/rutgers-lib/41978/.

Yu, Alan C. L. 2007. Understanding near mergers: The case of morphological tone in Cantonese. Phonology 24.187–214.

Yu, Dominic. 2000. The Underlying Representation of Cantonese Vowels. Stanford, CA: Stanford University unpublished undergraduate honors thesis.

Yue-Hashimoto, Anne. 1991. The Yue dialects. Languages and Dialects of China. Journal of Chinese Linguistics Monograph Series 3.294–324.

Yue-Hashimoto, Oi-kan. 1972. Phonology of Cantonese. Cambridge [England]: Cambridge University Press.

Zee, Eric. 1999. Chinese (Hong Kong Cantonese). Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet, 58–60. Cambridge: Cambridge University Press.

Zee, Eric. 2003. Frequency analysis of the vowels in Cantonese from 50 male and 50 female speakers. Proceedings of the 15th International Congress of Phonetic Sciences, 1117–1120. Universitat Autònoma de Barcelona Barcelona. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_1117.pdf.

Zhang, Zhenjiang. 2009. Language and society in early Hong Kong (1841-1884). Guangzhou, China: Zhongshan University Press.