Mergers-in-Progress in Hong Kong and Vancouver Cantonese-English Bilinguals

Lauretta Cheng

LING 449: Honours Thesis

Supervisor: Dr. Molly Babel

University of British Columbia
# Table of Contents

Abstract .......................................................................................................................... 3

1.0. Introduction ........................................................................................................... 4
  1.1. The HKC Mergers-in-Progress ................................................................. 5
  1.2. Internal and External Factors ................................................................. 7
  1.3. Immigrant and Homeland Populations ............................................... 12
  1.4. Research Question ...................................................................................... 15

2.0. Methods ............................................................................................................... 16
  2.1. Participants .................................................................
      2.1.1. Vancouver Participants ....................................................... 16
      2.1.2. Hong Kong Participants ..................................................... 17
  2.2. Materials ......................................................................................................... 18
      2.2.1. Stimuli .................................................................................. 18
      2.2.2. Language Questionnaire ...................................................... 19
      2.2.3. Post-Task Awareness Interview ............................................ 20
  2.3. Procedure ........................................................................................................ 21
      2.3.1. Production ........................................................................... 21
      2.3.2. Perception ............................................................................ 22

3.0. Results and Analysis ........................................................................................... 22

4.0. Discussion .......................................................................................................... 28
  4.1. [l]~[n] Continuum ....................................................................................... 28
  4.2. Ø~[ŋ] Continuum ....................................................................................... 29
  4.3. [m]~[ŋ̩] Continuum ................................................................................... 32
  4.4. General Discussion and Limitations ......................................................... 33

5.0. Conclusion .......................................................................................................... 37

Acknowledgments ....................................................................................................... 38

References .................................................................................................................. 39

Appendix .................................................................................................................... 43
The phenomenon referred to as 懶 音 laan5 jam1, or “lazy pronunciation”, in Hong Kong Cantonese (HKC) is a set of consonant mergers-in-progress that have been studied for many decades; these involve consonants in syllable-initial, syllable-final and syllabic positions (eg. Wong, 1941; Zee, 1999). A recent apparent time production study in Hong Kong reported that several of the syllable-initial mergers were nearing completion in the youngest generation (To, Mcleod & Cheung, 2015). While this sound change has been well-documented within Hong Kong, only a limited number of studies have examined Cantonese phonology in immigrant communities (e.g. Tse, 2016), and none appear to have targeted the consonant mergers. As such, the current study investigates both perception and production of a subset of the HKC mergers ([n-]→[l-], [ŋ-]↔Ø-, [ŋ]→[m]) in Vancouver’s immigrant Cantonese-speaking population, comparing across older and younger generations as well as to speakers in Hong Kong.

The perception experiment used a two-alternative forced-choice lexical identification task. Participants heard Cantonese words from 13-step minimal word-pair continua ranging from the innovative to conservative variant for each merger, and their task was to indicate which lexical item they heard. The production experiment was an isolated-word production task. Participants were prompted with both Chinese characters and the English translation to produce 22 Cantonese words containing the target contrasts. Finally, speaker awareness of the mergers was probed in a post-task interview, and bilingual dominance scores were calculated using the Bilingual Language Profile (Birdsong, Gertken & Amengual, 2012). The results of this study add a new perspective to scholarship on the HKC mergers and on the course of sound change in immigrant communities more generally, while also contributing to research on the phonetics and phonology of Cantonese heritage speakers using an experimental production and perception approach.
MERGERS-IN-PROGRESS IN CANTONESE-ENGLISH BILINGUALS

1.0. Introduction

Phonetic variation in Hong Kong Cantonese (HKC) consonants has been documented by numerous scholars over the last century, beginning in the 1940s (e.g. Wong, 1941; Zee, 1999; To, Mcleod & Cheung, 2015); this has led to a set of mergers-in-progress in syllable-initial, syllable-final and syllabic consonants. The Hong Kong community is somewhat aware of these ongoing changes, with the term 懶音 laam5 jaml “lazy pronunciation” being applied to the usage of innovative variants in colloquial speech (Chow, 2009). The current study focuses on three of the HKC consonant pairs – two syllable-initial and one syllabic – involved in the sound change: (1) [n-]~[l-] as in 男 naam4 ‘male’ and 藍 laam4 ‘blue’; (2) [ŋ]-Ø- (zero-initial) as in 牛 ngau4 ‘cow’ and 嘔 au2 ‘vomit’; and (3) [ŋ]-[m] as in 五 ng5 ‘five’ and 唔 m4 ‘not’ (see Table 1). We will expand on previous work by exploring the progression of the HKC mergers in the immigrant Cantonese-English bilingual population in Vancouver, Canada, as compared to bilinguals in Hong Kong. In doing so, we will also be the first to examine perception of the consonant mergers-in-progress, in addition to production, using an experimental approach.

<table>
<thead>
<tr>
<th>Merger</th>
<th>Chinese Character</th>
<th>Jyutping Romanization</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[n-]→[l-]</td>
<td>藍 laam4</td>
<td>naam4 laam4</td>
<td>‘blue’</td>
</tr>
<tr>
<td></td>
<td>男 naam4</td>
<td>laam4</td>
<td>‘male’</td>
</tr>
<tr>
<td>[ŋ]-Ø-</td>
<td>嘔 au2</td>
<td>ngau2</td>
<td>‘cow’</td>
</tr>
<tr>
<td></td>
<td>牛 ngau4</td>
<td>au4</td>
<td>‘vomit’</td>
</tr>
<tr>
<td>[ŋ]→[m]</td>
<td>唔 m4</td>
<td>ng5 m5</td>
<td>‘not’</td>
</tr>
<tr>
<td></td>
<td>五 ng5</td>
<td>m5</td>
<td>‘five’</td>
</tr>
</tbody>
</table>

Table 1. Examples of Cantonese words involved in each merger, and the direction of change.
1.1. The HKC Mergers-in-Progress

Early in the 20th century, all three sound pairs were considered clearly distinguished (Ball, 1907). Two pairs, /n-/~/l-/ and /ŋ̩-/~/m̩/, were phonemically contrastive; [ŋ̩-] and Ø-, on the other hand, were used in separate allophonic contexts where Ø- appeared before high tones, and [ŋ̩-] appeared before low tones (with a handful of colloquial forms proving exception to this). After some decades, Wong (1941) noted phonetic variation for two of these pairs, and Bauer (1983) later discovered the third ([ŋ̩]→[m̩]) in speakers born around 1936 to 1949. As such, the three mergers appeared to have initialized after the 1900s and before the 1940s. Table 2 details the documented timeline and trajectory of these sound changes.

While the [n-]→[l-] and [ŋ̩]→[m̩] mergers progressed steadily over time, [ŋ̩-] and Ø- showed evidence of merging in both directions. Yeung (1980) found that speakers born before 1960 replaced historical Ø- with [ŋ̩-] more than replacing historical [ŋ̩-] with Ø-, and at the same time speakers born after 1960 showed a reversed pattern. As the years passed, the phonetic variations also changed in how it was manifested. Bourgerie (1990) found in sociolinguistic interviews that use of the innovative variants were differentially governed by social and linguistic factors, including age, gender, formality of speech style and lexical category, with a mention of the potential effects of word frequency. The innovative [l-] and Ø- variants were more likely to be used younger and female speakers, as well as in less formal settings. Later, Zee (1999) concludes on the basis of past literature and articulatory data from a young Cantonese-speaking adult that (a) the historical initial /n-/ category had merged into [l-] in informal speech (with the potential of exceptions in isolated monosyllables, careful speech, or “hypercorrective” individuals) and (b) historical initial /ŋ̩-/ and Ø- categories had “split” into both [ŋ̩-] and Ø- (i.e., a change in both directions that erased the historical tone-based allophonic contrast).
Table 2. Summary of the diachronic progression of the three target mergers in Hong Kong including percentages of sound change occurrence found in each cohort, referenced from To et al. (2015).

More recently, To, Mcleod, and Cheung (2015) examined the progression of sound changes in Hong Kong by collecting isolated-word productions from 112 adults (aged 18-45) and 138 children (aged 10-12), all of whom were born and educated in Hong Kong, speaking HKC as their mother tongue. In a survey across the common Cantonese phonetic variations, they found that a few of the syllable-initial or syllabic variations were nearing completion, including the pairs of interest: [n]-[l-], [ŋ]-[Ø-], and [ŋ̩]-[m̩]. In terms of age stratification, [l-], Ø- and [m̩] are described as more likely to be used by younger speakers, while [ŋ]- was still more commonly used by older speakers. Other than boys producing more zero-initial than girls, no other case
showed significant gender differences. A caveat is that their results were based on elicitations of a single example for each sound, which may not necessarily generalize across all words in an individual’s lexicon and fails to capture within word variability within a single speaker. However, their findings were overall consistent with previous literature. In sum, the three HKC mergers-in-progress were initiated sometime in the early- to mid-1900s, and as of a few years ago in Hong Kong, appear to be have reached a stage where they are almost fully-merged. It should be noted that all findings in this review were based on production data, collected by various methods including palatography, isolated-word elicitations, and sociolinguistic interviews; however, while there are perception studies on the HKC tone mergers (e.g. Mok, Zuo & Wong, 2013), an extensive search did not uncover any studies investigating perception of the HKC consonant mergers.

1.2. Internal and External Factors

Both language-internal, such as phonetic interactions between sounds and contexts of a language, and language-external influences, such as social associations or language contact, may motivate sound change. Evidence suggests that linguistic changes are typically caused by the co-occurrence of multiple factors – both internal and external (Thomason, 2010); on the other hand, external social factors are also often involved in non-change or reversals (Hickey, 2012). In the context of the HKC initial consonant sound changes, areal linguistic research supports a mainly internal-based source considering that similar changes have also been documented in other Chinese varieties, such as in Guangzhou Cantonese, Southwestern Mandarin (Sichuan) and Jianghuai Mandarin (southern Jiangsu) (Ding, 2010). However, the multitude of political and social external factors surrounding Hong Kong and their Cantonese speakers also exert pressure
on the HKC mergers – in both directions. Motivations behind both advancement and resistance of the mergers are outlined in Table 3.

Language-internally, from the perspective of production, To, Mcleod and Cheung (2015) discuss how the mergers may stem from the “principle of articulatory ease”, where sounds move towards relatively simpler articulations. This could relate to the articulatory gestures involved in forming the single sound or due to co-articulation with nearby sounds frequently produced together. To et al. (2015) speculate that articulating [m] and Ø- requires less articulatory effort than the velar nasal [ŋ], in the case of [m] because both the tongue and lips would simply be in rest position while [ŋ] involves raising the tongue dorsum and lowering the jaw. A case of simplifying from co-articulation might be in common sequences of a syllabic velar nasal words followed by bilabials, such as ŋ ng5 ‘five’ in ng5 man → m5 man1 ‘five dollars’. A similar type of influence may originate from the visual salience of speech sounds, as proposed by Ohala (1986), where co-articulation of a sound sequence, and the lack of salient visual cues, particularly in the case of [ŋ] as compared to [m], may play a role in misinterpretation of sounds (see also McGuire and Babel, 2012). Additionally, due to the logographic nature of the Chinese orthography, which does not directly map symbols to sounds, and the lack of standardized romanization system, these sound changes cannot be impeded by literacy, though Whelpton (1999) does note that students will use Cantonese dictionaries that contain some type of Romanisation.

In terms of externally-motivated change, one major factor to consider is the contact between Cantonese and other Chinese dialects. Non-Hong Kong Cantonese and other Chinese varieties contain many of the phonetic variants that are entering HKC (To, Mcleod & Cheung, 2015). In addition, since the handover of Hong Kong in 1997 from British rule to Chinese, the Hong Kong
government has been promoting the “biliterate and trilingual” policy (兩文三語) – biliterate in Chinese and English while trilingual in Cantonese, Mandarin and English (Wang & Kirkpatrick, 2015). Though individual schools taught and used the two languages to different extents, Cantonese and English were already integrated into the education system; in 1998, Mandarin was added as a mandatory subject in the school curriculum of all primary and secondary schools. As such, the closely-related Mandarin variety of Chinese could play a role in the changes of the Cantonese sound system, particularly in younger generations who have been subject to this education system.

Mandarin has the clearly contrasted initial [l-] and [n-] phonemes, and the equivalent Cantonese words historically began with the same sounds, such as the Cantonese-Mandarin minimal pairs of ‘old’ (Cantonese lou5 vs. Mandarin lau3) and ‘brain’ (Cantonese nou5 vs. Mandarin nau3). Particularly with the pervasiveness of the Pinyin romanization system for Mandarin, many younger speakers are explicitly aware of the sounds and sound contrasts in Mandarin. While Bauer (1997) believed the influence of Mandarin would bolster HKC changes which increased similarity to Mandarin and attenuate changes that decreased similarity (e.g. [n-]→[l-]), Whelpton (1999) considered the possibility that local Hong Kong speakers may attempt to differentiate themselves from “mainlanders” (Chinese) in order to hold onto their identity as Hong Kongers even after the handover. In this scenario, HKC would continue on the path of current sound changes, even if that diverges from Mandarin, and Whelpton’s (1999) survey on social perceptions of different pronunciations suggested that Hong Kong residents did attach some significance to certain mergers, including [ŋ̩]→[m̩], for making local vs. outsider distinctions.
<table>
<thead>
<tr>
<th>Merger</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production</strong></td>
<td>Principle of articulatory ease: Preference for simpler articulations, as in [m] and Ø- Coarticulation: Assimilating to common sound sequences, especially to bilabials</td>
<td><strong>Language Contact</strong></td>
</tr>
<tr>
<td><strong>Perception</strong></td>
<td>Misperception: Lack of salient visual cues increase chance of misperception</td>
<td><strong>Social Importance</strong></td>
</tr>
<tr>
<td><strong>Orthography</strong></td>
<td>Literacy: Logograms do not directly map sounds</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reversal</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orthography</strong></td>
<td>Literacy: Logograms do not directly map sounds</td>
<td><strong>Language Contact</strong></td>
</tr>
<tr>
<td><strong>Prescriptivism</strong></td>
<td>Proper Pronunciation Campaigns: Promoting “proper” pronunciation especially in media, increasing public awareness and hypercorrection Cantonese Education: Teaching “correct” vs. “incorrect” pronunciations, framing sound changes negatively</td>
<td><strong>Speech Styles</strong></td>
</tr>
</tbody>
</table>

Table 3. Summary of language-internal and -external factors potentially involved in motivating either the continued mergers or a reversal in the HKC consonant sound changes.
In the education of Cantonese, there appears to be explicit direction of “correct” or “incorrect” pronunciations, particularly for the younger generations, raising awareness of the mergers in the students but also framing the innovative forms in a negative light. The rationale for this practice likely comes out of the campaign led by Hong Kong professor Richard Ho Man-Wui aimed at promoting “proper pronunciation” (正音 zing3jam1) of Cantonese, which has gained much attention and influence since the 1980s (Asprey, 2007). This perspective has especially impacted pronunciations in the media, and advanced the idea to the public of needing to prevent “lazy” ways of speaking. As a result of this prevailing prescriptivist approach, individuals may be subject to both conscious and unconscious influence to avoid these mergers and hypercorrect to what they believe is the conservative variant. In fact, these external factors in Hong Kong, and somewhat in Vancouver as well, seem posed to reverse or stall the progressing mergers, creating a “static situation where a change does not go to completion and where both incoming and outgoing forms… are possible” or where the conservative forms are still preserved in more formal registers (Hickey, 2012: 5).

Taking into account the prominent role of English in the experiences of many Cantonese speakers, some scholars have considered the influence of the English language on HKC. Most Hong Kong residents have at least some knowledge of English, though Cantonese is likely to be their dominant language. While English has been cited as having lexical, syntactic and intonational ramifications on HKC, including the well-known and pervasive use of English loan-words and code-mixing, the conclusion seems to be that English has only a minimal impact on the phonological system, if at all (Ding, 2010; Whelpton, 1999). On the other hand, it is uncertain whether the influence of English would become more pronounced given more input and higher levels proficiency, seeing as despite its overt prestige in Hong Kong, English is still
very much a secondary language in practice. In comparison, for HKC-speaking immigrants in Vancouver, English is the dominant language of their environment though they are still able to encounter Cantonese through family, community and media.

In terms of potential influences coming from English, the sound changes moving away from initial and syllabic [ŋ] would entail becoming more similar to the English sound system, which does not have [ŋ] in those syllable positions. On the contrary, since English has two clearly contrastive /n/ and /l/ phonemes, the obvious conclusion is that any effect of English should reverse this merger, in the same way as Mandarin, because of assimilation across phonological systems. However, native speaker descriptions of the Cantonese alveolar nasal /n/ suggest that it is phonetically different from English /n/, sounding “somewhere between English /l/ and /n/” or like a nasalized lateral. In light of this, it is also possible that experience with English phonemes could lead the bilingual speaker to dissimilate the two /n/ categories, and cause the Cantonese /n/ to further merge towards the /l/. If English language contact does play a measurable role in changes occurring in Cantonese, the amount of English experience may be a mediating factor in explaining differences in trajectory of the mergers.

1.3. Immigrant and Homeland Populations

While bilingual language contact has been shown to be a potential motivation for sound changes, it is important to take into account the fact that the populations we describe as “bilinguals” are not all the same even if we categorize them under one label. Because of this, it would appear to be more accurate to view bilinguals as existing on a continuum (Kagan & Polinsky, 2007). Within speakers of the same language pair, there can be extreme differences between the degree of proficiency of each language, or the language acquisition process of each individual. Bilinguals may have a stronger or weaker language, or they may be balanced; they
may live in a bilingual community where both languages can be used relatively equally, or they may mainly speak the majority language in the community and the minority language at home. These factors lead to differing language representations and degrees of language dominance.

With respect to Cantonese-English bilingual populations, individuals from Hong Kong and Vancouver generally have different language profiles in terms of language exposure and experience throughout their lives. Although individual language experiences within a city is by no means all the same, compared to Hong Kong-based bilinguals, Vancouver-based ones should be much more likely to use and be immersed in English. In particular, the group of Cantonese-English bilinguals who are the children of immigrants experience a vastly different linguistic environment during development than those who were raised in Hong Kong. These bilinguals – termed here as heritage speakers of Cantonese – grow up with English as the societally-dominant language used in their schools and community, and Cantonese as a minority language learned at home. Just like with the group of speakers labelled as “bilingual”, however, “heritage speakers” are a highly heterogeneous group where language experiences and levels of bilingual proficiency vary widely (Benmamoun, Montrul & Polinsky, 2013; Kupisch, 2013).

Linguistic transfer refers to a situation where aspects of one language become more similar to another due to language contact. At the phonetic level, previous work has demonstrated that a bilingual’s two languages can influence each other; more importantly, researchers have also found that the strength of transfer effects can vary across degree of proficiency, amount of usage, and age of language exposure (Flege, 1987; Flege, Schirru & MacKay, 2003; Chang, Yao, Haynes & Rhodes, 2011). In a comparison of the production English and Mandarin contrasts by L1, L2 and heritage bilingual speakers of Mandarin, Chang et al. (2011) found that heritage speakers were able to produce both within- and between-language
phonetic contrasts better than the late learners of either Mandarin or English. These results seem to indicate that the early exposure provided in heritage speakers’ experience leads to a different organization of their phonology where categories of similar sounds across languages are dissimilated, unlike the assimilated categories of late learners. In another example that explored the effects of bilingual dominance, early Spanish-Galician bilinguals were found to be worse at both perceiving and producing Galician mid vowel contrasts when they were Spanish-dominant but not when Galician-dominant (Amengual & Chamorro, 2015). That is, it appears that heritage phonological systems differ from L2 bilingual systems, and that being less dominant in a language can lead to more transfer from the dominant to non-dominant language.

Considering that bilingual contact-induced sound changes also originate from interactions between phonological systems, the evidence from bilingual phonetic transfer research suggests that differences in language history, input, and usage may lead to a distinction in terms of sound change progression between early (heritage) and late (L2 learner) bilinguals. In the case of HKC, this suggests that the merger may follow a different trajectory for Vancouver heritage Cantonese-English bilinguals, who are exposed to both languages early in life, compared to the bilinguals who were raised in Hong Kong and learned English as an L2 (i.e. both Hong Kong- and older Vancouver-based bilinguals). In fact, Kang and Nagy (2016) report just this situation in the Korean stop voice onset time (VOT) merger, where younger heritage speakers in Toronto patterned differently from older Toronto and Seoul Korean groups of speakers in terms of the merger. In this case, the authors propose that English dominance may be an explanation due to the importance of VOT for English stop contrasts, pointing out that the heritage speakers were the only group for whom English was the dominant language. With a similar set of language-external factors at play in the circumstances of Cantonese-English bilinguals in Hong Kong and
Vancouver, the patterns across age group and city could potentially show parallels to Kang and Nagy’s study of mergers in immigrant and homeland populations.

1.4. Research Question

The current study aims to describe the progression of the HKC mergers in Vancouver’s immigrant Cantonese-speaking population as compared to bilinguals in Hong Kong, with a particular interest in the Cantonese heritage speakers. Following the experimental production-perception approach to investigate sound change (e.g. Harrington, Kleber & Reubold, 2008; Bukmaier, Harrington & Kleber, 2014), this project explores the HKC mergers-in-progress through both perception and production domains. For the purposes of this thesis, I will be focusing on the results from the merger perception experiment.

We will examine the state of the mergers by comparing across speakers based on demographic grouping variables, including city, age group and gender. If the merger is progressing towards completion as the literature of HKC indicates, we would expect the younger generation of speakers to have sound categories that are relatively more merged than the older generation. As Hong Kong and Vancouver communities are geographically distant and have language environments unlike the other, the mergers may progress separately across cities, and based on the general trend in sound change research, we may expect gender differences especially where social status, networks and activities vary between males and females. In addition, younger Vancouver individuals may potentially demonstrate a difference in the merger progression from the other demographic groups due to their unique experience as early bilinguals growing up in an English-dominant location. One hypothesis is that bilinguals will follow the previously found trend in production and be more advanced in the HKC mergers if they are younger. If they are more advanced, we would expect to see less distinct categorization of the
two merging sounds, and if fully merged, we would expect to see no difference in perception of words starting with either of the two sounds. On the other hand, based on the plethora of external factors that seem to favour a stalling or reversal of the merger directions, an alternate hypothesis is that younger participants will be less merged in their perception of the categories. Given the lack of research specifically on the state of the consonant mergers within the domain of perception, this option appears to be equally plausible. Secondarily, owing to the differences found in heritage speaker phonological systems as well as sound changes occurring in other heritage language populations, we additionally hypothesize that Vancouver heritage speakers will pattern differently from the other speakers in either the extent or direction of merger. In short, a variety of outcomes are possible in the HKC perceptual mergers.

2.0. Methods

2.1. Participants

In Vancouver, the participants were recruited from the University of British Columbia as well as from the Greater Vancouver Area (henceforth referred to as “Vancouver”) community through public postings and snowball sampling. They were compensated either with course credit or payment ($15) for the one-and-a-half-hour session. In Hong Kong, participants were recruited from the Hong Kong Polytechnic University. At the time of writing, 89 participants have been recruited: 40 in Vancouver and 49 in Hong Kong. For each location, participants were further divided into two age groups, the older and younger generations.

2.1.1. Vancouver Participants

The criteria for older speakers were to have been born in Hong Kong, lived there until age 18, and have currently lived in Vancouver for around or over 20 years. This older group
consisted of 17 participants (11 F; 6 M) between age 52-64 with a mean age of 56. They had lived in Vancouver between 18-28 years, averaging 21 years across the group. A few participants were included who deviated slightly from these criteria. Two participants were born in Guangdong province in China before moving to Hong Kong at an early age. Another participant grew up in Hong Kong’s neighbouring region Macau, whose residents are considered to speak the same variety of Cantonese as is spoken in Hong Kong (i.e. HKC), before moving in Hong Kong at age 20.

The criteria for younger speakers were to have lived in Vancouver since age 5 or under, and have at least one parent who was an immigrant from Hong Kong. This younger Vancouver group consisted of 23 participants (12 F; 11 M) between age 17-24 with a mean age of 20. The majority were born in the Vancouver area, while one was born in another Canadian city and one born in Copenhagen, Denmark before moving to Vancouver at age 4 and 1, respectively. The remaining three were born in Hong Kong, but immigrated to Canada before age 3. The experiment was conducted either in a sound-attenuated booth at the University of British Columbia, or in a quiet classroom space at Kwantlen Polytechnic University.

2.1.2. Hong Kong Participants

The criteria for both older and younger speakers were that they had to have lived in Hong Kong since birth. The older group consisted of 21 participants (12 F; 9 M) between age 44-60 with a mean age of 54. The younger group consisted of 28 participants (13 F; 15 M) between age 17-27 with a mean age of 19. One younger male who had been born in the UK but had moved to Hong Kong before age 3 was also included.
2.2. Materials

2.2.1. Stimuli

Speech production stimuli consisted of 64 Cantonese words presented visually in the Chinese orthography accompanied by the English translation (refer to Appendix for full word list). Of these, 22 included the target sounds (4 each for [l-], [n-], [ŋ] and Ø- words; 3 each for [ŋ̩] and [m̩] words), while the remaining 42 were filler word pairs that contrasted non-target phonemes. To avoid missed trials from lower proficiency speakers (e.g. younger Vancouver speakers), the word list was designed to consist of minimal pairs and near-minimal pairs that would be familiar to heritage speakers with more limited vocabularies.

The speech perception stimuli consisted of 39 audio files from three 13-step synthesized speech continua. These speech continua were generated between Cantonese real-word minimal pairs that differed in only the target contrast. The minimal pairs are listed in Table 4.

<table>
<thead>
<tr>
<th>Target Merger</th>
<th>Jyutping Romanization</th>
<th>Chinese orthography</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[n]~[l]</td>
<td>lou5-nou5</td>
<td>老-脑</td>
<td>‘old’- ‘brain’</td>
</tr>
<tr>
<td>[ŋ]~Ø</td>
<td>aak1-ngaak1</td>
<td>握-呃</td>
<td>‘shake [hands]’- ‘deceive’</td>
</tr>
<tr>
<td>[ŋ̩]~[m̩]</td>
<td>m4-ng4</td>
<td>唔-吳</td>
<td>‘no, not’- ‘Ng (surname)’</td>
</tr>
</tbody>
</table>

Table 4. Minimal word pairs for each synthesized continuum.

First, a female native Cantonese speaker and trained linguist from Hong Kong was recorded producing the target minimal pairs using an AKG C 520 headset mic with a USBPre 2 Pre-Amp into Audacity 2.1.2 (2016) in a sound-attenuated booth. She read from transcriptions of each target word, accompanied by the Chinese orthography, to elicit the desired maximally contrastive pronunciations. The speaker was asked to produce each pair naturally but as similarly as she could aside from the target contrast to facilitate a more natural continuum synthesis. Each target word was produced at least 3 times. The clearest and most similar recorded pairs,
determined auditorily and through spectral analysis in Praat (Version 5.4.08, Boersma & Weenink, 2015) by features such as duration, pitch, and intonation, were selected for synthesis.

These six natural productions of Cantonese lexical items were then used as endpoints to create three word-pair continua using the STRAIGHT program in Matlab (Kawahara et al., 2008). The continua were created with 25 steps synthesized between the initial [l] and [n] (in lou4-nou4), the initial vowel [a] and [ŋ] (in aak1-ngaak1), and the syllabic [m] and [ŋ] (in m4-ng4). From these, every odd-numbered step was selected (i.e. 1, 3, 5, … 21, 23, 25) to result in 13 equidistant steps. Figure 1 illustrates the transition between the initial sound in one of the continua with waveforms and spectrograms of the mid- and endpoint recordings. Judgments from the author and Cantonese-speaking lab members were used to select the most natural sounding continua. In total, there were 39 tokens, 13 variants from each of the three continua.

![Waveforms and spectrograms of the endpoints and midpoint of the aak1-ngaak1 continuum.](image)

Figure 1. Waveforms and spectrograms of the endpoints and midpoint of the aak1-ngaak1 continuum. From left to right: (step 1) zero-initial [aːk̚], (step 13) maximally ambiguous initial, (step 25) velar nasal-intial [ŋaːk̚]

2.2.2. Language Questionnaire

The Bilingual Language Profile questionnaire, originally designed by Birdsong, Gertken and Amengual (2012), is a publicly available resource that calculates a bilingual dominance score between two languages of a bilingual speaker using self-report responses to questions pertaining to four sections – language history, usage, proficiency and attitudes. The online
version hosted on Google Forms automatically tabulates the scores. For the purposes of this study, we adapted the online questionnaire into two versions to compare English and Cantonese: one for Vancouver participants and one for Hong Kong participants. Both were administered in English. A small number of changes were made to better reflect the language experiences of the Cantonese-English bilingual and heritage population. The modifications included clarifying questions relating to education due to the different types of education systems available in Hong Kong and Vancouver, and to account for Mandarin experience alongside Cantonese and English. In addition, general questions about the individuals’ biographical information, language background and language experiences were included into one questionnaire.

In terms of scoring, each section is tabulated, then weighted equally to calculate a score between -218 to +218, where 0 indicates balanced dominance between the two languages, -218 indicates complete Cantonese dominance and +218 indicates complete English dominance. In other words, higher absolute scores suggest asymmetrical dominance in either direction while positive and negative scores signal relative English and Cantonese dominance, respectively.

2.2.3. Post-Task Awareness Interview

A post-experiment interview was conducted to probe linguistic awareness about the mergers in participants’ own speech and experiences. Experimenters in both Vancouver and Hong Kong were trained to follow the same interview protocol. The experimenter conducted the interview in English, supplemented by Cantonese (in Vancouver) or Mandarin (in Hong Kong) when necessary to ensure comprehension. Participants were presented with a list of minimal or near-minimal word pairs that were designed to contrast historically on the target sounds and asked to say each word in pairs, as they would in casual speech. For each pair, they were asked to explain any differences between their productions and to name the prescribed or “correct”
pronunciation if they were aware of any. To determine whether explicit knowledge of “correct” forms were interfering with perceptual judgements during the experiment, they were asked about their strategy for responding during the perception task. At the end, participants were asked if they were familiar with the term 懶音 "lazy pronunciation", what they thought it meant and whether they believe, or have been told by others, that they use it in their own speech.

2.3. Procedure

After being seated in a sound-attenuated booth or quiet classroom space, participants first completed the production task, immediately followed by the perception task. After this, participants were asked to fill out the language questionnaire and take part in the exit interview.

2.3.1. Production

In the production experiment, participants were prompted visually with both Chinese characters and the English translation to say 64 isolated Cantonese words. The words were presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) in three randomized blocks, where each word appeared only once per block for a total of 192 tokens. They were asked to read the Chinese characters out loud, but if they did not know how to read it, to translate the English gloss into Cantonese and say the word out loud. If they were translating, they were told they could take as much time as they needed to think, and if they could think of multiple possible translations, to use the most common spoken variant that they knew. The task was self-paced, where participants pressed a button on the keyboard to begin the two-second recording for each trial and could re-record if they wished before moving on to the next word. Most participants took about 20-30 minutes for this portion of the experiment, though some younger Vancouver speakers took up to 45 minutes to complete the task. Productions were recorded using an AKG C 520 headset microphone with a USBPre 2 Pre-Amp into E-Prime.
2.3.2. Perception

The perception task was a two-alternative forced choice lexical identification task, where participants were instructed to respond as quickly as possible and to not overthink their response. Audio stimuli, the synthesized Cantonese words, were presented over AKG K 240 Studio Semi-Open Stereo headphones at a comfortable listening level using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA), accompanied by a visual display of the appropriate minimal pair Cantonese words in Chinese orthography (characters) and the English translation. Each trial consisted of one audio file playing as the computer screen presented two choices: one word on the left labelled with a ‘1’ (corresponding to the response on the button box), and the other word on the right labelled with ‘5’. For example, if the participant hears a word on the lou4-nou4 continuum, they would see 老 ‘old’ (lou4) labelled with ‘1’ and 腦 ‘brain’ (nou4) labelled with ‘5’. Participants were given 3 seconds after the end of the audio file to respond on the button box (either ‘1’ or ‘5’). If no response was registered within 3 seconds, the next trial would begin automatically. There were three blocks of randomized trials. Each token was repeated 3 times throughout the experiment, for a total of 117 trials (3 continua x 13 steps x 3 repetitions).

3.0. Results and Analysis

Analyses for the perception data were conducted in R (R Core Team, 2015). As a confirmation of the language dominance composition of the groups as divided by City and Age (Hong Kong or Vancouver, older or younger), the four were compared on the calculated English-Cantonese bilingual dominance scores, depicted in Figure 2. As expected, older Hong Kong participants [M= -82.0, SE=9.07], younger Hong Kong participants [M= -87.23, SE=4.86], as well
as older Vancouver participants [M=-94.82, SE=6.48] had negative dominance scores, indicating clear Cantonese dominance for these three groups of individuals who had grown up in Hong Kong. In comparison, younger Vancouver participants, the only group to have grown up outside of Hong Kong speaking Cantonese as a heritage language, demonstrated English dominance with positive dominance scores [M=78.69, SE=6.48].

Figure 2. Distribution of language dominance scores from Cantonese- (negative) to English-dominant (positive) for the four groups divided by City and Age.

Turning to the effect of group factors on the response patterns in the perception task, trials with null responses were first removed from the data set, amounting to 1.7% of the data. Three binomial generalized linear mixed effects models, one for each continuum ([l]~[n], Ø~[ɲ], and [m]~[ŋ]), were constructed in R (2015) using the lme4 package (Bates, Maechler, Bolker & Walker, 2015). The dependent variable was the proportion of responses for the lexical item with the novel variant (i.e. lou5, aak1, m4) and the fixed categorical factors were City (Hong Kong, Vancouver), Age Group (Old, Young) and Gender (Female, Male). Each included a random
intercept of Subject and random slope of Step. Due to initial model convergence, the “bobyqa”
Optimizer was also added to each model to solve this issue.

For the [l-]~[n-] contrast, there were main effects of Step (β = -0.29, SE = 0.11, z = -2.78, p<0.01) and Age (β = 2.58, SE = 0.11, z = -2.78, p<0.01), as well as a significant interaction between Step and Age (β = -0.43, SE = 0.15, z = -2.80, p<0.01); no effects were found for City or Gender. Across all participants, younger listeners are more categorical in their responses on
the two ends of the perceptual continuum than older listeners (see Figure 3).

![Figure 3](image)

Figure 3. Proportion of responses for the innovative /l/ variant at each step of the [l-]~[n-] continuum (from 1-13) between older and younger age groups.

For the Ø-~[ŋ-] contrast, there was no main effect of Step, which means that averaging across all participants, the initial Ø- and [ŋ-] words were not differentiated significantly by their onset pronunciation. However, there was both a significant two-way interaction between Age
MERGERS-IN-PROGRESS IN CANTONESE-ENGLISH BILINGUALS

and City ($\beta = 2.69$, SE = 1.04, $z = -2.58$, $p<0.01$), and a three-way interaction of Step, Age and City ($\beta = -0.44$, SE = 0.15, $z = -2.86$, $p<0.01$). The interaction of Age and City stems from younger Vancouver participants being more likely to respond with the conservative [ŋ]- variant overall [$M=0.45$, SE=0.02], differing from Hong Kong participants who leaned towards responding with the innovative variant (both older [$M=0.53$, SE=0.03] and younger [$M=0.54$, SE=0.03]). As seen in Figure 4, the step interaction shows that younger Hong Kong and older Vancouver listeners were more categorical than older Hong Kong and younger Vancouver listeners across the continuum; the two cities display opposite effects of age. Notably, when listeners were responding more categorically, the direction was in fact reverse to the historical pattern, with listeners judging pronunciations on the [ŋ]- end of the continuum as the word that historically had the zero-initial variant, and vice versa for the $\emptyset$- end of the continuum.

![/∅/-/ŋ/ Continuum by Age and City](image)

Figure 4. Proportion of responses for the null variant at each step of the $\emptyset$-~[ŋ]- continuum (from 1-13) between older and younger age groups in female and male listeners.
MERGERS-IN-PROGRESS IN CANTONESE-ENGLISH BILINGUALS

For the [m]~[ŋ] contrast, main effects of Age ($\beta = 2.45$, SE = 0.72, $z = 3.43$, p<0.001) and Gender ($\beta = 1.63$, SE = 0.78, $z = 2.08$, p<0.05) were found. Although there was again no main effect of Step, there were significant two-way interactions between Step and Age ($\beta = -0.23$, SE = 0.08, $z = -2.83$, p<0.01) as well as Age and Gender ($\beta = -3.73$, SE = 1.04, $z = -3.40$, p<0.001). Moreover, significant three-way interactions included those between Step, Age and Gender ($\beta = 0.31$, SE = 0.12, $z = 2.60$, p<0.01) as well as City, Age, and Gender ($\beta = 3.98$, SE = 1.59, $z = 2.51$, p<0.05).

Figure 5. Proportion of responses for the innovative /m/ variant at each step of the [m]~[ŋ] continuum (from 1-13) between older and younger age groups in female and male listeners.

Younger participants appear to be slightly more categorical than older participants, but as seen in Figure 5, the three-way interaction of Step, Age and Gender shows that it is only the
younger females that are much more categorical than males and older females, while older and younger males did not differ markedly.

Overall, older females respond with the conservative [ŋ] word more than the males and younger females; however, this is again qualified by a three-way interaction, with City this time, which shows that only older Hong Kong females [M=0.48, SE=0.06] are more conservative than younger Hong Kong females [M=0.62, SE=0.04], while older Vancouver females are roughly equivalent in their proportion of responses to younger Vancouver females. On the other hand, older Hong Kong males and younger Vancouver males pattern together as opposed to younger Hong Kong males and older Vancouver males. These relationships are depicted in Figure 6.

![Figure 6. Proportion of total responses for the innovative-variant word on the [m]~[ŋ] continuum by male and female speakers of older and younger age groups and in both Vancouver and Hong Kong.](image-url)
4.0. Discussion

In this study, Cantonese-English bilinguals from both Hong Kong and Vancouver participated in production and perception tasks involving three HKC initial consonant mergers. As our study constitutes a first look at perception in these specific mergers, a variety of scenarios are possible. The perception task was a binary forced-choice lexical identification task, designed to examine how participants perceived a continuum of pronunciations for minimal pairs with critical merging sounds. The three target mergers ([l]~[n], Õ-~[ŋ], [m]~[ŋ]) have been flagged as nearing completion in the speech production of a group of children within Hong Kong (To, Mcleod & Cheung, 2015). The production results from To et al. (2015) are a particularly interesting comparison for the currently presented data, as their sample of Hong Kong children (at the time of the study, aged 10-11) consisted of speakers born in 1992-1993, which is of the same cohort as our younger speaker group, the majority of whom were born between 1992-1998 (at the time of study, aged 18-24, aside from one participant at age 27). The present study finds, in the domain of speech perception, age group effects in all three continua, each with varying interactions with other group factors. This suggests that each pair is in the process of merging perceptually but are, compared to each other, at different stages or subject to different external factors.

4.1. [l]~[n] Continuum

Younger listeners were generally more categorical in their responses to the [l]~[n] continuum, demonstrating less merged perceptual categories than older listeners regardless of city or gender. If younger participants are further along in the merger, we would expect the opposite age pattern, suggesting that the sound change in this sample is going in the reverse direction and unmerging instead. Despite the general trend found in the HKC literature thus far,
in the context of the Hong Kong and Vancouver’s social and political landscape, an unmerging of [l]~[n] is not an entirely surprising finding. In fact, there many converging external factors that would encourage a reversal of the merger-in-progress, or at least impede its progression. On the language contact front, the increased Mandarin and English contact in the experiences of the younger generation in both cities could be contributing to the reversal, as both languages contain clear phonemic distinctions between [l] and [n]. Further reinforcing the influence from the other languages is the fact that both Mandarin and English use roman letters to represent sounds (in the Pinyin Romanization system for Mandarin), which could serve to increase phonological awareness, and therefore awareness in variation of pronunciation in Cantonese as well. Additionally, the efforts to “correct lazy pronunciations” from the media and education system, along with the ability to easily identify and label sounds as “l” or “n” (as opposed to labelling the Ø-~[ŋ] pair which don’t have direct sound-letter correspondence), make this merger pair an easy target for hypercorrection. Even if the participants themselves do not hypercorrect, the heightened societal awareness in recent years may have led to more exposure to the conservative form through media and formal speech, and with the other factors at play, this may explain why younger speakers, as compared to older speakers, are more likely to keep the phonemes distinct.

4.2. Ø-~[ŋ-] Continuum

The results for the Ø-~[ŋ-] contrast showed that responses were divided between both variants across all continuum steps; on the whole, whether participants heard words that began with a vowel, velar nasal or an ambiguous sound in between, they were about equally likely to think they heard the Cantonese word for ‘shake [hands]’ 握 as they were to think they heard the word for ‘deceive’ 呃. This apparent lack of lexical differentiation based on the onset suggests that this merger is quite far along in its progression towards completion. However, not all
demographic groups performed in the same way. Rather than having a flat pattern across the continuum, the responses of younger Hong Kong and older Vancouver listeners instead showed indications of some more categorical responses. Interestingly, the resulting categorization reveals that these listeners were more likely to select the historically [ŋ]-initial word when hearing zero-initial tokens, and the historically zero-initial word when hearing [ŋ]-initial tokens, demonstrating an apparent category reversal.

One interpretation of the results draw on the fact that the sound change has gone in both directions over the years, first [ŋ-] → Ø- then Ø- →[ŋ-]. From a diachronic perspective, it is clear that the changes occurring in this contrast have been rather unstable: speakers born in the first half of the 1900s were found to merge exclusively towards [ŋ]; around the 1960s, a shift occurred in which speakers began to merge towards zero-initial and exhibited much less of the Ø- →[ŋ-] change (Yeung, 1980). Following that, nearly two decades ago, Zee (1999: 156) posited that “both the historical syllable-initial [ŋ] and historical zero-initial have split into the [ŋ]-initial and zero-initial” and that at the time, there was a higher proportion of zero-initial productions. It may be possible that rather than a split, changes in both direction are occurring simultaneously, maintaining a distinction but reversing the categories. If this were true, it may stem from the extra awareness and hypercorrection surrounding the mergers brought on by prescriptivism, compounded with the confusion of historical forms due to the variation across the past century and loss of the allophonic contrast.

However, before accepting this explanation, the effect of the specific minimal word-pair must be considered. Our experimental paradigm required exact minimal pairs to be used for creation of synthesized word continua. Since the historical tonal complementary distribution of Ø- and [ŋ-] ensured that there typically did not exist minimal pairs for this contrast, the one we
used included an exception to the rule where, reportedly, 呃 ngaak1 ‘deceive’ was the exception as it had a [ŋ] onset paired with a high tone (Bourgerie, 1990). Yet without the allophonic rule to refer to, it is not immediately clear which form was the historical version in actuality; as such, the onset of this word pair may have been switched in modern times, potentially with prescriptivist support. Under this scenario, the results would still suggest unmerging for this contrast if generalized, but in the more typically expected direction back to historical forms.

In context of the findings for the other two continua, a merger reversal in younger Hong Kong listeners with influence from prescribed forms in their education and media exposure would be a consistent account. On the other hand, what may be the source of increased categorization in the case of the older Vancouver group, who differ from the same age group in Hong Kong? One proposal might appeal to the effect of living in the Vancouver immigrant community – the network size of Cantonese speakers is far more restricted compared to in Hong Kong, leading to less exposure to variability; at the same time, Hong Kong Cantonese media, such as news programs and radio, are still widely available, potentially increasing the proportion of input from prescriptivist forms. Another possibility is that the discrepancy may originate from pre-existing group-level characteristics, such as education, status and other social factors, differentiating those who immigrated and those who did not, i.e. older Vancouver and older Hong Kong participants. While the collected background information has not yet been inspected, individuals who had the means to immigrate likely comprise to some extent a homogenous group with relatively high levels of education and socioeconomic status, whereas the Hong Kong group could be more heterogenous on these factors. Of potential significance in interpreting this finding is that our samples of older listeners did differ across cities, such as a smaller sample size in
Vancouver (n=17, as compared to n=21 in Hong Kong) as well as a relatively older and more restricted age range (the Vancouver sample had no participants under 52 but three above age 60).

4.3. [m]~[ŋ] Continuum

In the [m]~[ŋ] contrast, younger females across both cities are much more categorical in their lexical identification responses than any other demographic, while the other groups appear to show only slight trends for perceiving differences across this continuum. One way to understand this could be to see it as, like the [l]~[n] situation, a reversal in the merger, possibly under the same influences. Following the gender trends found in the sound change literature, young females leading this change would be a reasonable proposition. In terms of overall endorsement of each variant, older Hong Kong females were the only group to choose the [ŋ] word more often than [m] at 52% of the time, which contrasts with younger Hong Kong females who chose the [m] word 62% of the time. In addition, older Hong Kong and younger Vancouver males selected the [m] word relatively often at around 66% of the time. This may suggest that these groups are merging towards the innovative variant. In comparison, Vancouver females, older Vancouver males and younger Hong Kong males all hover just above 50% in their selection of the [m] word, seeming quite evenly merged. Generally, we also find that older males had higher variability in their responses than the other groups. This could potentially reflect differences in social networks, activities and status between genders and age groups. However, caution must be applied to any interpretation of the data.

A large caveat to the results for the [m]~[ŋ] continuum is that the minimal pair used differed very much in terms of their distribution and usage, one being a very common negation word and the other a surname, which is written in English as the letters “Ng”. Firstly, the English orthography of the surname may cue some to the historical pronunciation leading to inflated
categorization, though post-experiment interviews revealed that many participants did not seem to be aware that the spelling related to the pronunciation. The special status of a surname, as opposed to the typical content word may also affect representations and the rate of change involving it. Secondly, any skew in responses towards the innovative variant of [m] may be driven by the higher frequency and therefore faster accessibility of the word for ‘not’ 唔 m4, in comparison to the surname. Thirdly, some of the younger Vancouver respondents noted in their post-experiment interview that they did not actually know the word for the surname “Ng” in Cantonese, while others mentioned that even if they might recognize it, they’d never had to use it in the context of speaking Cantonese in Vancouver. As a consequence, the results particularly from young Vancouver speakers may not be a reliable indicator of the state of the merger in perception. In fact, the younger Vancouver male bias for responding with the innovative [m] could potentially be attributed to the knowledge gap. The rationale behind using this minimal pair was because it was already the best available choice: the number of syllabic words are extremely limited, to the extent that 唔 m4 ‘not’ is the single [m] Cantonese word, and out of the three homonyms for ng4, the surname was the most common. This does limit the conclusions we can draw from the perception results; however, the production data contains elicitations of more common [ŋ] words, such as 五 ng5 ‘five’, which may help to understand the state of this merger.

4.4. General Discussion and Limitations

In the findings from this perception experiment, the [n-]→[l-], [ŋ-]→Ø and [ŋ]→[m] HKC consonant mergers-in-progress appear to be in perceptual flux rather than having reached a state of completion. The recurring age effect across merger contrasts point in the direction of there being at least a subset of younger participants who show more categorical responses along each continuum. For [l]~[n] it was all younger participants, for Ø-~[ŋ-] it was younger Hong
Kong listeners (along with older Vancouver listeners), and for [m]-[ŋ] it was younger females. Judging from the largely indiscriminate response patterns across the continua in the other demographic groups, many HKC speakers didn’t differentiate the minimal pairs based on onset pronunciations, an indication of a full merger on a perceptual level. With members of the younger generation behaving differently than the older, these data seem to suggest an age-related trend of reversals for mergers which were previously quite far along in the process. As such, the evidence does not support the hypothesis that younger speakers would be more merged, which was based on the historical trajectory found in change in production. Further, in no case are heritage speakers acting uniquely, leaving that hypothesis unsupported as well. While the alternative hypothesis of a reversal is backed by these perception data, it is unclear whether such a result is due to a discrepancy between change in perception as compared to production, or whether trends in production are also undergoing a reversal, which would be a new finding.

The overarching theme across these merger perception findings seems to be the language-external influences surrounding HKC and how those have changed over recent years to impact younger generations, namely increased language contact and enforcement of prescriptivism. Younger speakers across Hong Kong and Vancouver do not appear to differ drastically in their perception patterns other than the [ŋ]- and zero-initial pair, which is intriguing in and of itself. There are many differences between these two groups, including their language proficiency, exposure and usage, as encapsulated in their contrasting dominance scores. One might expect stronger influences in Hong Kong youth from (i) prescriptivism through their schooling and exposure to Hong Kong media, (ii) exposure to more variation in Cantonese speech (including in formal registers), as well as (iii) contact with Mandarin due to its compulsory instruction in Hong
Kong schools. It would be interesting to further dissect more fine-grained predictors of the perception patterns, as well as compare performance on production across groups.

In general, interpretations of the current results are limited by constraints related to the experimental paradigm, in which the task lacked control for confounds in the stimulus items and may have introduced a bias for making more categorical distinctions. This perception experiment did not account for various word-level factors, such as frequency or word-class, and social factors, such as speech register, which have previously been found to play a role in the HKC sound changes (Bourgerie, 1990). The words used in the experiment were not all equally common due to constraints on the number of minimal pairs with these contrasts. For the same reason, only one perceptual continuum was created for each merger. As it is, the [l]~[n] pair are relatively frequent words, the Ø-~[ŋ] pair less frequent, and the [m]~[ŋ] pair differing with one extremely frequent function word and the other a much less common surname, as discussed previously. Of these three minimal pairs, the latter two were the only options available for those contrasts, even without accounting for the need to select words likely known by heritage speakers. The consequence is that the word frequency is a confound, as words of differing in frequency undergo sound change at different rates. Accordingly, the results may potentially be reflective only of perception for our specific words and word pairs without being generalizable to perception of the mergers as a whole. Acknowledging this, the varying levels of word frequency for our chosen words may explain some of the by-continuum differences found in our results.

While scholars have noted the effect of speech register on the pronunciation of words involved in the HKC mergers-in-progress (Bourgerie, 1990; Zee, 1999), we were unable to control for the effect of register and styles in our experiment. It is likely the case that participants who are more aware use a more formal register in production due to being in an unnatural
setting, and they may draw on their explicit knowledge of “correct” or formal pronunciations in the perception task. Other limitations include our sample size, which is especially small with regards to the difficult-to-recruit older male participants. Considering several social factors (city, age group, gender) in our analysis further reduces the power of inference from these data. Future plans include continued data collection. As well, demographic groups may have been unbalanced on various social dimensions, leading to confounds that underlie differences in performance; for example, there appears to be higher education level and English ability in older males than older females in Vancouver. Further work may include factor analysis of these extraneous determinants to uncover specific predictors of the mergers-in-progress.

Although the findings from this perception experiment might seem to run counter to the most recent To et al. (2015) overview in Hong Kong that suggest more merging in younger participants, it is possible that the perception of this merger has been proceeding with a different trajectory than has production. For instance, the mergers may be advancing towards full completion in the informal speech of the younger generation, while their perceptual representations are accommodating the variability in speech that they hear from the older generation, media and more formal speech; in this way, their lexical representations may encode phonological forms for both informal and formal registers. Due to the demands of the lexical decision task which forces listeners to make distinctions between minimal pairs, these individuals may draw on their knowledge of formal pronunciations to chose a response, despite not typically producing them this way. If this were the case, it would align with Whelpton’s (1999: 55) prediction that regardless of all the external impediments to sound changes in formal registers of Cantonese, “informal speech will continue to go its own way” and “the gap between the ends of the stylistic continuum will remain particularly wide” in HKC.
On the other hand, the results from To et al. (2015) for the three mergers in question were based on productions of a single, frequent word for each target pair; this implies that the values from their study may overestimate the rate of change occurring even in informal speech production as it cannot be generalized to entire phonological categories. Whether the state of the HKC mergers in perception and production are progressing in tandem or not, forthcoming analysis of production data from the current study will help to shed light on the remaining questions.

5.0. Conclusion

This study provides a first investigation into the perception of Hong Kong Cantonese consonant mergers-in-progress in both Vancouver and Hong Kong bilingual listeners. Each of the three mergers we examined presented evidence of an age effect where a subset of younger speakers exhibited more categorical perception, potentially suggestive of merger reversals. Performance by heritage Cantonese speakers from Vancouver do not appear to differ significantly from the other demographic groups. Contrary to our findings in the perceptual domain, existing HKC sound change literature on speech production have indicated continued merging towards completion in younger generations; based on current results, the source of divergence from these previous findings is still unclear. Future directions include analysing the accompanying production data, comparing between perception and production results, as well as investigating more thoroughly the specific predictors that lead to group differences and individual variation.
Acknowledgements

This research was partially funded by the UBC Language Sciences Initiative. I’d like to thank our collaborators Yao Yao and Chang Liu who ran the experiment at the Hong Kong Polytechnic University, as well as Dr. Heather Bliss and Paivi Koskinen at the Kwantlen Polytechnic University for their contributions. Last but not least, thank you to all the members of the Speech in Context Lab, particularly Sophie Bishop, Cassandra Savage and Shannon Briggs, and my supervisor Molly Babel for all of their help, guidance and advice throughout this project.
References


Wong, S. L. (1941). A Chinese syllabary pronounced according to the Cantonese dialect. Hong Kong: Chung Hwa Book Co.


## Appendix

### Production Word List

<table>
<thead>
<tr>
<th>Chinese Character</th>
<th>English Gloss</th>
<th>Jyutping Romanization</th>
<th>Chinese Character</th>
<th>English Gloss</th>
<th>Jyutping Romanization</th>
</tr>
</thead>
<tbody>
<tr>
<td>老</td>
<td>old (eg. person)</td>
<td>lou5</td>
<td>跟</td>
<td>to follow</td>
<td>gan1</td>
</tr>
<tr>
<td>腦</td>
<td>brain</td>
<td>nou5</td>
<td>羹</td>
<td>spoon</td>
<td>gang1</td>
</tr>
<tr>
<td>藍</td>
<td>blue</td>
<td>laam4</td>
<td>乾</td>
<td>dry</td>
<td>gon1</td>
</tr>
<tr>
<td>男</td>
<td>male; man</td>
<td>naam4</td>
<td>八</td>
<td>eight</td>
<td>baat1</td>
</tr>
<tr>
<td>濕</td>
<td>[rain] jacket; coat</td>
<td>lau1</td>
<td>百</td>
<td>[one] hundred</td>
<td>baak1</td>
</tr>
<tr>
<td>慄</td>
<td>to be angry</td>
<td>nau1</td>
<td>滑</td>
<td>smooth</td>
<td>waat6</td>
</tr>
<tr>
<td>龍</td>
<td>dragon</td>
<td>lung4</td>
<td>畫</td>
<td>to draw</td>
<td>waak6</td>
</tr>
<tr>
<td>農(夫)</td>
<td>farmer</td>
<td>nung4</td>
<td>(口)渴</td>
<td>thirsty</td>
<td>hot3</td>
</tr>
<tr>
<td>握(手)</td>
<td>to shake [hands]</td>
<td>aak1</td>
<td>殼</td>
<td>shell</td>
<td>hok3</td>
</tr>
<tr>
<td>呃</td>
<td>to deceive</td>
<td>ngaak1</td>
<td>釀</td>
<td>ugly; disgraceful</td>
<td>cau2</td>
</tr>
<tr>
<td>嘔</td>
<td>to vomit</td>
<td>au2</td>
<td>酒</td>
<td>wine; alcohol</td>
<td>zou2</td>
</tr>
<tr>
<td>牛</td>
<td>cow</td>
<td>ngau4</td>
<td>草</td>
<td>grass</td>
<td>cou2</td>
</tr>
<tr>
<td>愛</td>
<td>love</td>
<td>oi3</td>
<td>早</td>
<td>early</td>
<td>zou2</td>
</tr>
<tr>
<td>我</td>
<td>I, me</td>
<td>ngo5</td>
<td>車</td>
<td>car</td>
<td>ce1</td>
</tr>
<tr>
<td>矮</td>
<td>short (eg. person)</td>
<td>ai2</td>
<td>姐</td>
<td>older sister</td>
<td>ze1</td>
</tr>
<tr>
<td>危(險)</td>
<td>dangerous</td>
<td>ngai4</td>
<td>跑</td>
<td>to go running</td>
<td>paaau2</td>
</tr>
<tr>
<td>唔</td>
<td>no, not (eg. good)</td>
<td>m4</td>
<td>飽</td>
<td>to be full</td>
<td>baau2</td>
</tr>
<tr>
<td>五</td>
<td>five</td>
<td>ng5</td>
<td>趴</td>
<td>to lie on stomach</td>
<td>paa1</td>
</tr>
<tr>
<td>唔該</td>
<td>please; thank you</td>
<td>m4 goi1</td>
<td>巴(士)</td>
<td>bus</td>
<td>baa1</td>
</tr>
<tr>
<td>誤會</td>
<td>misunderstanding</td>
<td>ng6 wui6</td>
<td>拍(手)</td>
<td>to clap [hands]</td>
<td>paa1</td>
</tr>
<tr>
<td>唔同</td>
<td>different</td>
<td>m4 tung4</td>
<td>免</td>
<td>rabbit</td>
<td>tou3</td>
</tr>
<tr>
<td>午餐</td>
<td>lunch</td>
<td>ng5 caan1</td>
<td>刀</td>
<td>knife</td>
<td>dou1</td>
</tr>
<tr>
<td>呢(個)</td>
<td>this one</td>
<td>ni1</td>
<td>太(陽)</td>
<td>sun</td>
<td>taa13</td>
</tr>
<tr>
<td>佢</td>
<td>s/he; him/her</td>
<td>keoi5</td>
<td>帶</td>
<td>to bring</td>
<td>daai3</td>
</tr>
<tr>
<td>國(家)</td>
<td>country; nation</td>
<td>gwok3</td>
<td>跳</td>
<td>to jump</td>
<td>tiu3</td>
</tr>
<tr>
<td>角</td>
<td>corner</td>
<td>gok3</td>
<td>釣(魚)</td>
<td>to fish; fishing</td>
<td>diu3</td>
</tr>
<tr>
<td>廣(東話)</td>
<td>Cantonese</td>
<td>gwong2</td>
<td>咳</td>
<td>to cough</td>
<td>kat1</td>
</tr>
<tr>
<td>講</td>
<td>to say; speak</td>
<td>gong2</td>
<td>金</td>
<td>gold</td>
<td>gam1</td>
</tr>
<tr>
<td>果(汁)</td>
<td>fruit juice</td>
<td>gwo2</td>
<td>橋</td>
<td>bridge</td>
<td>kiu4</td>
</tr>
<tr>
<td>嘅(個)</td>
<td>that one</td>
<td>go2</td>
<td>叫</td>
<td>to call; be known as</td>
<td>giu4</td>
</tr>
<tr>
<td>山</td>
<td>mountain</td>
<td>saan1</td>
<td>起</td>
<td>to stand</td>
<td>kei5</td>
</tr>
<tr>
<td>生</td>
<td>raw; to be born</td>
<td>saang1</td>
<td>記</td>
<td>to remember</td>
<td>gei3</td>
</tr>
</tbody>
</table>