PERCEPTUAL IDENTIFICATION OF TALKER ETHNICITY IN VANCOUVER ENGLISH

by

Phoebe Wong

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Bachelor of Arts

in

The Faculty of Arts
(Honours in Linguistics)

University of British Columbia
(Vancouver)

Supervisor: Dr. Molly Babel

April 2014

© Phoebe Wong, 2014
Abstract

As a result of historical and ongoing immigration patterns, English in Metro Vancouver has been in close contact with Punjabi and Cantonese for years. This persistent language contact, the establishment of Chinese and East Indian communities, and the struggle of Canadian-raised Punjabi and Chinese identities set up ripe conditions for the formation of Chinese and East Indian ethnolinguistic repertoires. Thus, the purpose of this project was to investigate listeners’ ability to identify the ethnicities of white, East Indian, and Chinese talkers in Vancouver in order to determine the existence of such repertoires. Semi-spontaneous speech of talkers from these three groups was recorded. Then listeners from Metro Vancouver were presented with sentences from these recordings and asked to identify the talker’s ethnicity for each sentence in a forced-choice identification task. Results indicate that listeners were able to identify talker ethnicity above chance, implying the emergence of salient ethnolinguistically-associated features contributing to the diversity of Vancouver English. In addition, listeners with Chinese social networks exhibited higher accuracy identifying Chinese talkers than listeners with white social networks, emphasizing the role of experience in perceiving talker ethnicity. Finally, results indicating that listeners were most likely to misidentify East Indian and Chinese talkers as white as opposed to the other Asian talker group argue against the existence of a general Asian ethnolinguistic repertoire. These results also suggest that Cantonese and Punjabi Canadian communities can preserve elements of their linguistic heritage while establishing a space in the mainstream Vancouver community.
# Table of Contents

Abstract ii  
Table of Contents iii  
List of Tables v  
List of Figures vi  
Acknowledgements vii  
1 Introduction 1  
1.1 Theoretical background 4  
1.2 Historical background 14  
1.3 Research problem 23  
2 Methodology 25  
2.1 Talking 25  
2.1.1 Talkers 25  
2.1.2 Procedure 27  
2.1.2.1. Collection of read speech 28  
2.1.2.2. Collection of semi-spontaneous speech 28  
2.2 Listening 29  
2.2.1 Listeners 29  
2.2.2 Stimuli 29  
2.2.3 Procedure 31  
3 Results 33  
4 Discussion 40  
4.1 Findings 40  
4.2 Potential limitations 43  
4.3 Future Directions 48  
5 Conclusion 56  
References 57  
Appendix A: The North Wind and the Sun 64
Appendix B: Talker Background Questionnaire 65
Appendix C: Listener Background Questionnaire 67
List of Tables

Table 1: Listener mean identification accuracy rates 33
Table 2: Regression model 1 35
Table 3: Regression model 2 36
Table 4: Regression model 3 36
Table 5: Confusion matrix data 38
List of Figures

Figure 1: The Pearl River Delta region in Guangdong 16
Figure 2: Listener mean talker identification accuracy rates 34
Figure 3: Listener identification accuracy by network rank 37
Figure 4: Confusion matrix 38
Figure 5: Bronfenbrenner’s bioecological model 50
Acknowledgements

First of all I would like to thank Dr. Molly Babel, my supervisor, for her continual knowledge, guidance, and patience as we worked on this project. I would also like to thank my family for their support and encouragement, and Drs. Stefan Dollinger and Anita Szakay for helping inspire this project. In addition, I would like to thank Jenny Wang for her insight on and explanations of the Mandarin phrases and words discussed here; Michael McAuliffe, for his technical support and general advice; and Jamie Russell, for her accompaniment, suggestions, and humor.
1 Introduction

Since its establishment in 1886, the city of Vancouver, British Columbia, has been multilingual and multicultural, with British-ancestry Anglophones intermingling with First Nations, Francophones, and more recent immigrants from around the world. Today, the greater region of Metro Vancouver still retains some of its diverse characteristics. In particular, while English is still the dominant language of the region, Asian languages of are in close competition for linguistic space, specifically Cantonese and Punjabi.

Assuming that English is the dominant language of Metro Vancouver is based on two main justifications. The first is prestige. English continues to assert its dominance through federal, provincial, and municipal institutions as well as the media. English is one of the two official languages of Canada (the other one being French) and the de facto official language of British Columbia (BC), the province in which Vancouver resides. In addition, in the BC Language Education Policy states that all primary and secondary public school students are expected to achieve proficiency in English (British Columbia Ministry of Education, 2004). Primary and secondary independent (non-public) schools must include English Language Arts in their curriculum (British Columbia Ministry of Education, n.d.).

The second justification for assuming that English is the dominant language of Metro Vancouver is that it is the language of the majority of the region’s population. About fifty-nine percent of Metro Vancouver respondents on the 2011 Census listed English as a mother tongue, with ninety-six percent of that subgroup of respondents, or fifty-two percent of all respondents, having only English as their mother tongue (Statistics Canada, 2012b). Furthermore, of the subgroup of respondents listing English as a mother tongue, ninety-six percent did not list another mother tongue. In other words, the majority of those with English L1s did not have another L1, and those with exclusively English L1s are estimated to make up fifty-six percent of Metro Vancouver’s population. In addition, eighty percent of Metro Vancouver respondents on the same census listed English as one of the languages they most frequently used at home. Again, the vast majority (ninety-two percent) of these responses did not include languages other
than English, so it can be assumed that the majority (seventy-two percent) of respondents reported exclusively uses English at home. Lastly, the most frequent reported ethnic origin on the 2011 National Household Survey was English, comprising twenty percent of all responses, and third most common was Scottish, comprising fifteen percent of all respondents (Statistics Canada, 2014c).

On the other hand, French, Canada’s other official language, is only reported as a home language or mother tongue by a paltry proportion of respondents. Including both single and multiple responses, not even two percent of Metro Vancouver respondents listed French as a mother tongue, and less than one percent listed French as their most frequent home language (Statistics Canada, 2012b). Therefore, despite its status as an official language federally, French appears to have only a marginal presence in Metro Vancouver with respect to language usage and knowledge.

In contrast, immigrant languages appear to occupy much more of the non-institutionalized linguistic space of Vancouver. Approximately forty-one percent of the 2011 Metro Vancouver census single responses named non-official non-Aboriginal languages as their mother tongue (Statistics Canada, 2012b), which matches the proportion of Vancouver’s population that is comprised of foreign-born immigrants (Statistics Canada, 2013a). For the most frequently used home language, non-Aboriginal non-official languages constituted about twenty-seven percent of single responses on the 2011 census (Statistics Canada, 2012b). Unfortunately, multiple responses could not be considered, as the Canadian census does not appear to differentiate between Aboriginal and non-Aboriginal non-official languages when considering multiple responses.

In addition, an estimated forty-five percent of Metro Vancouver’s population consists of visible minorities (Statistics Canada, 2013a). Here I follow the Canadian Employment Equity Act’s definition of visible minority, which describes members of visible minorities as “persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour” (Statistics Canada, 2012c). Thus, while white people still comprise an absolute majority in Metro Vancouver and can be considered dominant in that sense (as well as in terms of power, historically, which will be discussed in the next section), they are still only a slight majority.
In particular, Chinese and South Asians are the listed as the two largest visible minority groups in Metro Vancouver by the Canadian National Household survey (Statistics Canada, 2014c), with Chinese comprising eighteen percent of Metro Vancouver’s population, and South Asians comprising eleven percent. Furthermore, Chinese was the second most commonly reported ethnic origin after English on the 2011 National Household Survey (Statistics Canada, 2014c).

Linguistically speaking, the two most frequent single responses after English for both mother tongue and most often used home language were Punjabi and Cantonese on the 2011 Census (Statistics Canada, 2012b). Approximately six percent of single respondents listed Punjabi as their mother tongue, as well as their most frequent home language. Cantonese was reported as the exclusive mother tongue and most frequent home language, in six percent and five percent of single responses, respectively.

As mentioned previously, these two languages and their associated ethnic groups have been in close contact with Anglophone whites in Metro Vancouver since the period in which the city of Vancouver was founded. This has occurred through fluctuating but prolonged immigration of Cantonese Chinese from predominantly Hong Kong and the province of Guangdong in the People’s Republic of China (PRC), and Punjabi East Indians primarily from the province of Punjab in India. Thus, the formation and maintenance of such communities and the continual language contact between English, Punjabi, and Cantonese comprise the motivation for the current paper.

This continual language contact leads one to imagine that perhaps these Asian languages have influenced English in some way, or perhaps vice-versa, such that even Canadian-raised Punjabis and Cantonese may use linguistic features in English derived from their heritage languages. In this way, Asian immigrant groups in Metro Vancouver may contribute to English diversity in the same way that global ‘World Englishes’ have done in countries where non-Anglophone heritage language speakers constitute the population majority (e.g. Pennycook, 2011).

In addition, it was speculated here that English usage might diverge from the dominant variety, particularly in the descendants of ethnolinguistic immigrant groups who are visible minorities. That is, the ability to visibly distinguish that talkers are from different ethnic backgrounds and the resulting consequences, such as discrimination and
potentially subsequent isolation of visible minority groups may serve to reinforce differences between dominant and immigrant descendant varieties of English. Indeed, Statistics Canada (2013a) reports that Chinese and South Asians constitute the top two visible minority groups in Metro Vancouver.

Thus this project endeavors to answer the question: does the English of native or near-native Metro Vancouver-raised talkers differ between ethnic groups, specifically among Punjabi East Indian, monolingual Anglophone white, and Cantonese Chinese talkers? Since language may differ between groups in many linguistic domains and in many features within those domains, it was decided to use people who should be experts at detecting such differences. In other words, can Vancouver-raised listeners identify the ethnicity of talkers from these three ethnic groups? If the answer is yes, it can be inferred that differences must exist which serve as perceptual cues distinguishing talkers from these three groups.

1.1 Theoretical background

First of all, it is necessary to address and define certain terminology. All definitions apply to the terms used throughout the paper, including previous usage, unless they are described otherwise in specific sections. It is also important to keep in mind that the definitions of such terms are generally operational in nature and sometimes unfortunately ad-hoc. I do not necessarily always endorse and justify conceptualizing these terms in these manners on a theoretical level or using them outside of this paper.

To start off, the traditional distinction between language and dialect is problematic for many reasons. Perhaps the most major reason involves the erroneous perceptions associated with using the traditional distinction between ‘languages’ and ‘dialects.’ That is, distinguishing between languages and dialects can imply a rigidly categorical hierarchy that is empirically unjustified and ill-defined. Traditional theories have proposed that languages are mutually unintelligible, while dialects of a language are mutually intelligible, but this is not always the case (Chambers & Trudgill, 2009). For example, the majority of people in both Inner Mongolia, PRC, and in the nation of
Mongolia both speak the Mongolian ‘language.’ However, the influence of Chinese on
Mongolian spoken in Inner Mongolia and, on the other hand, Russian on Mongolian
spoken in Mongolia is cited as one factor leading to frequent mutual lack of intelligibility
between Mongolian talkers from the two different regions (Mühlhäusler, Lee-Smith, &
Wurm, 1996). Thus it has been argued that the term ‘language variety,’ in not drawing a
rigid line between dialect and language, avoids some of these misconceptions.

However, ‘language variety’ may still be overly categorical and discrete in nature
for the purposes of describing ethnolinguistic variation, which is emergent. This
assumption is erroneous in the face of evidence that talkers may use features associated
with a given ‘ethnic language variety’ sometimes, and features associated with another
one at other times. Alternatively, Benor (2010) has proposed the concept of an
ethnolinguistic repertoire, which she defines as “a fluid set of linguistic resources that
members of an ethnic group may use variably as they index their ethnic identities” (p.
160). I find this definition suitable to describe any potential features characterizing the
speech associated with different ethnic groups in Vancouver in general, albeit a few
modifications.

First of all, it may be better to say that these linguistic resources are not ones that
ethnic group members may “use variably as they index their ethnic identities,” but those
that ethnic group members “may use variably as their ethnic identities are indexed.”
Substituting the active voice for the passive one is probably more appropriate here, as the
active voice may imply that talkers are the only agents in projecting their ethnic identity.
Although this is certainly an empowering statement, unfortunately this may not always be
the case. I argue that people’s ethnolinguistic identities are not only self-constructions but
also projected onto them by greater society, and the latter may in turn shape self-
identifications of ethnicity. Benor (2010) does emphasize this point, but fails to
incorporate it in her formal definition. As well, if language transfer really does underlie
potential variation between ethnic groups, it may well begin before Cantonese and
Punjabi Canadians even begin talking. It is hard to attribute agency if the mechanisms of
transfer are ones over which the talkers may have little control.

In addition, I find it is necessary here to make a distinction between languages,
language varieties, and ethnolinguistic repertoires within them. This is not an entirely
hierarchical or categorical relationship, and it must be emphasized that, like ethnicity and race, language usage is probably better described as probabilistic and gradient. However, the use of the term ethnolinguistic repertoire to describe any type of language use is, I argue, overly ambiguous.

Principally, I borrow the notions from various approaches to language classification and creolistics, as mentioned in Noonan (2010), to distinguish between Punjabi, Cantonese, and English as different languages, while ethnolinguistic repertoires can be nested as variations within a given language. Firstly, under more traditional family tree models, these three languages belong to different language families and subfamilies. Cantonese would be classified under the Yue branch of Sinitic languages from the Sino-Tibetan language family. Punjabi and English both would be classified as languages in the Indo-European family, but Punjabi is considered Indo-Aryan whereas English is considered Indo-European, specifically–Germanic (Lewis, Simons, & Fennig, 2014). Secondly, creolistic scholars may nest creoles as subdivisions under their proportionally dominant lexifier languages, if there is one. This is not to say that I am subsuming ethnolinguistic repertoires as creoles, especially since doing so might be implying some sort of speciation that may be unwarranted. However, using this criteria it may be useful to see Cantonese Canadian and Punjabi Canadian ethnolinguistic repertoires, if they exist, as still belonging to the general English language as opposed to Cantonese or Punjabi. This is especially appropriate given that talker elicitation materials here were in English, as well as in order to stress that potential differences in English speech between ethnic groups are viewed here as facilitating the diversification of English and not enforcing prejudicial stereotypes.

In addition, Benor (2010) also mentions that ethnolinguistic repertoires are not necessarily only used by members of the ethnic group they are associated with. Taking this with the two other aforementioned points, I define the term ethnolinguistic repertoire, for my purposes, as a fluid set of linguistic resources within a language that talkers may use variably as their ethnic identities are indexed. In addition, I will continue to use the term language variety to refer to what many scholars see as subdivisions of Yue Chinese (e.g. Guangzhou Cantonese and Toisan), English (such as Californian English and Canadian English) and Punjabi (e.g. Mahji or Malwa). I do not use
ethnolinguistic repertoire to describe these subdivisions, as it does not seem apparent that different ethnicities are associated with these different subdivisions.

I also must define here is my usage of ethnic/racial labels in this paper. One issue with this is that Punjabi and Chinese may be seen as traditionally referring to ethnicities, and white, on the other hand, as referring to race. However, as with the terms ‘language’ and ‘dialect,’ it is not clear to me where exactly to draw the line between race and ethnicity, and it seems that consensus within academia is also lacking (Chandra, 2012). Seeing that this is the case, I persist in the use of these labels and may use the terms ‘race’ and ‘ethnicity’ somewhat interchangeably, due to the blurred lines of this distinction.

Another valid criticism would be that at times, I tend to use the terms ‘Cantonese’ and ‘Chinese’ interchangeably, as well as ‘Punjabi,’ ‘East Indian,’ and ‘South Asian’ here. Although it may not be justified, I will clarify my use of such terms now. As a side note, I ask the question as to whether any set of terms, no matter how carefully justified, will not be criticized at least one person, no matter how carefully chosen, just given how complex the constructs of race and ethnicity really are, though of course some terms are far more offensive than others.

In the general discussion of this paper, I would like to make a distinction between these ethnic/racial terms. When using the term ‘South Asian,’ I refer to all those with ancestral ties to the Indian subcontinent. When using the term, ‘East Indian,’ I am referring to all people whose recent ancestors are comprised of those who would be Indian nationals today. Both ‘South Asian’ and ‘East Indian’ include Punjabis, but not only Punjabis. The term ‘Punjabi’ is used here to denote an ethnolinguistic group that can be seen as a subdivision as East Indian or South Asian, with ancestral roots in the region of Punjab in modern day India and Pakistan. Nayar (2012) describes ‘Punjabi’ as a term used to refer to people from the region of Punjab in present-day India and Pakistan. Nayar (2012) states that “the term Punjab (five waters) was first used by the Persians to refer to the geographic region of the northwestern part of the Indian subcontinent, where five rivers (Jhelum, Chenab, Ravi, Beas, and Sutlej) merge into the Indus River” (p. 5).

In the general discussion of this paper, ‘Chinese’ used to refer to anyone who self-identifies as such, and/or those with ancestral roots in the current day Han Chinese-
dominated parts of the PRC, Hong Kong, and Macau. ‘Cantonese’ is used here cautiously
to refer to all language varieties of Yue. This is somewhat problematic, given that some
speakers of Yue varieties other than the standard Cantonese varieties localized in
Guangzhou and Hong Kong, or those that are bilingual in both standard Cantonese and
other Yue varieties, consider their variety as distinct from Cantonese. For example, my
father and some of my classmates, who speak the Toisanese (Toisan-hwa) variety
of Yue Chinese but are not fluent in standard Cantonese, consider themselves speakers of
Toisan, but not Cantonese. As well, Toisan talkers often face discrimination and stigma
when speaking Toisan, by speakers of the standard varieties Cantonese (Leung, 2012).
However, at the same time, many speakers of Toisan would simply tell people that they
speak Cantonese to avoid confusion, long explanations, and perhaps condescension when
telling them that they do not actually speak Cantonese, but Toisan. Additionally,
perception of Cantonese as separate from other Yue language varieties seems to vary
between speakers (Leung, 2012). This suggests that talkers in our study who reported
speaking ‘Cantonese’ may include speakers of Yue language varieties other than the
standard versions of Cantonese, and this is the principal motivation between this
operational definition of ‘Cantonese.’ However, I would like to remind the reader that
this definition is merely operational, and I do not endorse the view that ‘Cantonese’
should be seen as an all-encompassing category for Yue language varieties. In other
contexts, it would probably be more appropriate to typify specific versions of Cantonese
as language varieties themselves under the Yue branch of Chinese to avoid privileging
standard versions of Cantonese over marginalized Yue varieties such as Toisan.

‘White’ is an extremely contentious term and difficult to define. One might
wonder why I would even want to include ‘white’ talkers in the study if the goal was to
look at Punjabi and Cantonese Canadian talkers. However, for one thing, those who are
self- and other-identified as ‘white’ simply do not suffer the same degree of
discrimination as those who are visible minorities. As well, I argue that, because of less
visual differences, even whites who are first generation immigrants are likely to be less
discriminated against and therefore less segregated. In turn, first generation immigrant
whites are therefore likely mix into the dominant Anglo-Canadian ‘mainstream’ more,
generally speaking. Thus, using the term ‘white’ provides us with a succinct and useful
reference level to measure the potential effects of ethnolinguistically associated change, as also mentioned by Benor (2010). That is not to say that the repertoires of visible minorities are inferior, but rather that points of comparison can be useful. For these reasons, I will use *white* to describe people who are consistently self- and other-identified as ‘white.’

Briefly, I will also clarify my use of the term ‘Canadian’ to refer primarily to those who were raised in Canada since age eight or younger. This is done for simplicity, though I recognize that immigrants who have arrived later in life may also consider themselves Canadian, and I wholeheartedly support their right to do so. I will also use of the word ‘Vancouver’ to refer to Metro Vancouver, or Vancouver CMA (Census Metropolitan Area), unless the city of Vancouver is expressly mentioned.

When describing the talkers who participated in this study, however, ‘East Indian’ and ‘Chinese’ refer to those whose ancestral language backgrounds are Cantonese and Punjabi. ‘White’ talkers refer to monolingual (except one French-English bilingual) Anglophones who are presumably, but not necessarily, of largely European ancestry. (By monolingual I am referring to most frequently used home language.) This is done because the labels presented to listeners for the Cantonese and Punjabi heritage language talkers were ‘Chinese,’ and ‘East Indian,’ respectively, as it was feared that not all listeners would be familiar with terms as specific as ‘Cantonese’ or ‘Punjabi.’ In addition, it must be stressed that these talkers confirmed their ethnolinguistic backgrounds, as elucidated by preliminary eligibility checks and post-participation background questionnaires, further described in the methods section.

Now I can go on to explain the conditions that make ripe the possibility of Chinese Canadian and East Indian Canadian ethnolinguistic repertoires in Metro Vancouver. One of these is, as mentioned previously, the existence of language contact. Kerswill (2010) outlines several potential outcomes of language contact, which include, for example, creolization and koinezation. The potential outcome focused on here is the use of heritage-language associated features in the dominant language of the locale, particularly by L1 speakers of those heritage languages, possibly leading to distinct ethnolinguistic English repertoires.
One mechanism by which this might happen is language transfer. I use a modified definition from Hickey (2010) and Sankoff (2008) to define language transfer as the transference of features from one language into another, resulting in (relatively) long-term and stable language shift. The term ‘transfer’ is used as opposed to ‘interference’ to characterize the influence of one language on another (Sankoff, 2008). This is because ‘interference’ may imply that language mixing in bilinguals and L2 language learners may be due to failure to keep languages separate, which may not necessarily be the case (Cutler, 2012).

According to Sankoff (2008), language contact consequences are determined in part by history of social relations among populations, including economic, political, and demographic factors. She says that rapid assimilation tends to be the norm except in cases of relatively stable bilingual communities. In can be argued here that the Punjabi and Cantonese communities in Metro Vancouver constitute such communities. While proficiency in their heritage language may vary between individuals in these communities, the ongoing influx of first generation Punjabi- and Cantonese-dominant immigrants continues re-introducing L1 Punjabi and Cantonese into these communities (Statistics Canada, 2014c), thereby stabilizing bilingualism in the community, though perhaps not necessarily in each individual member. Sankoff (2008) and Thomason (2010) reinforce this notion, by proposing that the degree of assimilation to the dominant language is negatively correlated with length of language contact.

Sankoff (2008) also argues that the influence of immigrant languages on local language tends to be very limited, except in cases where immigrants’ descendants have become dominant with respect to either numbers or prestige in the larger community. It is unclear whether the Cantonese and Punjabi Canadian communities of Metro Vancouver can be characterized in such a way. Obviously, white Anglophone culture is dominant in the global sense of Metro Vancouver, when considering overall numeric majority, historical power relations, and institutionalized emphasis.

However, Punjabi and Cantonese do appear to be the next most common languages in the area, despite the fact that their absolute proportions appear quite small when looking at single responses from the census. In addition, the fact that no further detail was given for multiple responses that included non-official languages hinders the
ability to make reliable estimates of the true extent of the usage and exposure to these languages in Metro Vancouver. As well, Cantonese and Punjabi do appear to be more common in certain municipalities in Metro Vancouver rather than the region as a whole. For example, the 2011 census for the city of Richmond reports about sixteen and fifteen percent of single respondents listing Cantonese as their mother tongue and most frequent home language, respectively (Statistics Canada, 2014a). The 2011 census for the city of Surrey reports that approximately twenty-one and seventeen percent of single respondents listed Punjabi as their mother tongue and most frequent home language, respectively (Statistics Canada, 2014b).

As far as the prestige of such languages goes, it does appear that these languages are dominant in some regards, as evidenced by the frequency of Punjabi and Cantonese signage in Metro Vancouver, as well as the existence of business and media outlets catering to Punjabi and Cantonese language audiences (Nayar, 2004; Ng, 1999; Yee, 2006). Furthermore, it is possible that Punjabi and Cantonese may be more prestigious at a more local scale, particularly in areas with high proportions of South Asians and Chinese, such as Richmond and Surrey. In any case, it is certainly true that Punjabi and Cantonese Canadians in the area have fought hard to create a linguistic space for their heritage language.

One reason that heritage-language derived traits may continue to be used even by L1 talkers of the dominant language is to “mark affiliation or to inscribe new social boundaries” (Sharma & Sankaran, 2011, p. 402). In this way, even though second-generation immigrants may not speak their heritage language fluently, using heritage-language derived traits may be a way for them to still preserve their cultural heritage identity while speaking English (Leung, 2012). Alternatively, the identity establishment motivation for formation of an ethnolinguistic repertoire may not be a result of language contact and/or transfer, per se. Canadian-raised immigrants face some unique challenges that can be quite different from those faced by white Canadians or immigrants who came to Canada as adults or during adolescence. Canadian-raised Punjabis and Chinese continue to struggle for their identity, as they are not necessarily accepted by Chinese and Punjabi immigrants who were raised abroad, or the dominant white Canadian society.
The former group includes not only their peers who immigrated after childhood, but also their older relatives who were raised in their ancestral region.

This struggle for identity can already be evidenced in lexical usage regarding these two groups, particularly among words and phrases used to refer to South Asians and Chinese who are raised abroad. For example, Ng (1999) notes that during the postwar period, Chinese-raised elderly first generation immigrants have referred to younger Canadian-raised Chinese as hànjìān (Mandarin)/hon3 gaan1 (Cantonese), meaning ‘traitor to the Han people,’ due to being overly xīhuà (Mandarin)/sai1 fa3 (Cantonese), meaning ‘Westernized’. One of my Mandarin-speaking friends who first came here recently as a university student from mainland China says that both of these terms are used today to refer to Canadian-raised Chinese by newcomers from the mainland such as herself. She describes the usage of the former when referring to Canadian-raised Chinese as primarily derogatory, and employed mostly in the context of tension between newly arrived Chinese immigrants and Canadian-raised Chinese. However, the valence she attributes to xīhuà is neutral.

Today, the terms coconut and banana are used to refer to ‘whitewashed’ South and East Asians, respectively, by both in-group and out-group members. It is said that like bananas, ‘whitewashed’ East Asians are yellow on the outside but white on the inside (Yee, 2006), and ‘whitewashed’ South Asians are brown on the outside but white on the inside. There exist reports documenting negatively valenced usage of coconut and banana. For example, in 1978, a UBC student residence newsletter published an article that used humor perceived by some as offensive to Asians (Yee, 2006). When someone replied to the article with commentary somewhat apologetic to the original authors, they were quickly faced with angry counter-replies, including one that accused the apologetic commentator of having a “banana attitude” (p. 176).

Nayar (2012) describes another example, reported by a Punjabi Canadian who moved from Prince Rupert, BC, a city where around three percent of the residents identify as South Asian (Statistics Canada, 2013b) to Metro Toronto, where around fifteen percent of residents report being South Asian (Statistics Canada, 2013c). Nayar (2012) quotes this individual as saying, “Some of the [Metro Toronto] Punjabi kids jokingly called me a ‘coconut’ or ‘Oreo.’ Although they were Punjabi and clowning
around, I didn’t like being called ‘white on the inside’ because I had spent my childhood sticking up for my brownness” (p. 246). On the other hand, the Chinese Varsity Club often uses a banana logo on signs advertising the club, reflecting a positively valenced in-group appropriation of ‘banana’ as well as the club’s overall affiliation with Canadian-raised Chinese (Ho, 2012).

In addition, terms have been established specifically to refer to Canadian-raised Chinese and South Asians, which fluctuate between terms and within terms with respect to associated valences. I have observed that the term ‘CBC’, standing for Canadian-born Chinese, continues to be used regularly by Canadian-born Chinese themselves as a way to distance themselves from first generation immigrants from Hong Kong and mainland China, whom Canadian-born Chinese may consider to have poor manners and English skills. On the other hand, first generation Chinese who immigrated after childhood may use ‘CBC’ in a derogatory way to reflect their impression that Canadian-born Chinese have shamed their families by abandoning Chinese culture (Yee, 2006). I also use the term CBC in this paper, but in a neutral way to refer to Chinese born in Canada.

A parallel Chinese term, tūshēng huáyì (Mandarin)/tou2 saang1 waa6 jeoi6 (Cantonese), or simply tūshēng (Mandarin) /tou2 saang1 (Cantonese), appears to have existed to refer to Canadian-raised Chinese, and more generally Chinese raised outside of China. The full phrase roughly translates as ‘local-born Chinese.’ Ng (1999) appears to use tūshēng in a neutral way to refer to Canadian-raised Chinese. On the other hand, as quoted by Ng (1999), Lee (1967) seems to use tūshēng to refer to the same group negatively, whom he appears to consider culturally deficient and suffering from an inferiority complex precisely due to their “poor command of the Chinese language” (p. 188-189). One of the Canadian-born Cantonese-speaking listeners of the present study said that the former part of phrase, tou2 saang1, is today extremely common, and neutral in usage. She also said that she had not heard the full term tou2 saang1 waa6 jeoi6. On the other hand, while the same newly arrived friend of mine I mentioned earlier said that the full Mandarin phrase tūshēng huáyì was not very frequent, although she did report hearing the phrase at least once before. Furthermore, she said that she had not heard the former part of the phrase used, tūshēng, but only the latter part, huáyì, which was quite common. She described huáyì as neutral in valence, but the full phrase as negative.
Most importantly, the principle function of language is communication. Whether you are an African elephant nose fish of the species *gnathonemus petersii* sending electrical pulses to another African elephant nose fish (Ahlborn, 2004), or a human speaking English to another human, the point is to get the message across to another conspecific. Thus using heritage-language or English is ultimately a function of whether these will be effective in conveying your message to your conspecific. There is less motivation to use a language if it risks compromised intelligibility or social isolation. Thus immigrants and their descendants are pressured to use English in order to communicate in government institutional settings, such as school and law courts. On the other hand, staying in touch with first generation immigrant parents or grandparents may require knowledge of the heritage language or adaptation of English to make sure one is understood as well as to maintain ties with one’s family. In this way, the roles of talker accommodation and face-to-face transfer of variants within a speech community are emphasized (Thomason, 2010).

In summary, the presence of language contact, prolonged and ongoing immigration, the formation of ethnolinguistic communities, and the more general necessity to communicate with others introduces the possibility that East Indians and Chinese in Vancouver may have ethnolinguistic repertoires in English differing from whites.

### 1.2 Historical background

Before 1961, ninety percent of Canadians had been born in Europe, and just three percent in Asia (Sandercock & Attili, 2009). Nonetheless, Chinese people have been in British Columbia since the 1850s, and Punjabis from at least the 1890s (Nayar, 2012; Ng, 1999). The first Chinese came to British Columbia from San Francisco in 1858. Two years later more sailed directly from China (Yee, 2006).

Most of the Chinese in Vancouver emigrated from Guangdong in the early period of Vancouver history, particularly the Pearl River Delta region of Siyi, which includes the
‘four counties’ of Taishan (variantly Toisan), Xinhui, Kaiping, and Enping (Ng, 1999; Yee, 2006). A map of the Pearl River delta region is shown below in Figure 1.

Emigration from Guangdong had been spurred by two particular previously occurring phenomena. Firstly, Guangdong’s population almost doubled between 1787 and 1850, but no agricultural solutions arose to increase food supply. Secondly, the First Opium War between China and Great Britain during 1839 to 1842 decreased the availability of certain jobs, and peasant tenants faced rising rents due to poor land ownership distribution (Yee, 2006).

In the decades before the official establishment of the city of Vancouver, Chinese were already beginning to settle along the Fraser River in the 1860s and 1870s (Ng, 1999). By 1884, one hundred fourteen Chinese were reported living on the Burrard Inlet. In 1885, the Canadian Pacific Railway announced the extension of its western terminus into the ‘mud flat’ north of False Creek, which would later become Vancouver’s Chinatown, or the area of the largest concentration of Chinese in Canada by the first decade of the twentieth century (Ng, 1999).

During the same year of 1885, one year before the city of Vancouver’s founding in 1886, Canada attempted to limit Chinese immigration by introducing the Chinese Immigration Act (Yee, 2006), one of several examples of specifically Chinese-targeted exclusionary federal policy. This act required all new Chinese immigrants to pay a head tax of fifty Canadian dollars, which in those days was worth considerably more than it is today. In addition, no inbound ship was permitted to carry more than one Chinese person per fifty tons. While this act was associated with a large immigration dip for a few years, by 1890, the rate of Chinese immigration went up again as Chinese families scrambled together enough money to pay the head tax. In addition, Chinese immigrants during the end of the 1800s and beginning of the 1900s were often refused citizenship (Yee, 2006).

In the early 1900s, the Chinese population comprised about one fifth of the Vancouver municipality’s total (Sandbeck & Attili, 2009). However, the both institutionalized and popular racial discrimination ensured the maintenance of an Anglo-European dominant culture (Sandbeck & Attili, 2009). Meanwhile, Punjabi Sikhs established the first Sikh organization, the Khalsa Diwan Society, in 1906 (Nayar, 2004). Sikh is the religion of many Punjabis.
An important event during this period was the occurred on September 7, 1907, when Vancouver’s two-thousand-member Asiatic Exclusion League staged a parade calling for an end to Asian immigration (Yee, 2006). This parade quickly degenerated into a riot, as members of the league raged through Vancouver’s Chinatown and Japantown, damaging both Chinese and Japanese businesses. This event is known today as the 1907 Vancouver Anti-Asian Riot, and was followed by the Chinese holding a three-day citywide general strike.

Figure 1: The Pearl River Delta region in Guangdong, present-day People’s Republic of China. Taken from Ng (1999).

The Canadian government’s response to this event was somewhat paradoxical. In 1908, Chinatown merchants were awarded financial reparations for the damage incurred and the business lost (Yee, 2006). On the other hand, earlier that year, another legal
barrier against Asian immigration was erected. The Continuous Journey Clause required all international immigrant-carrying ships arriving to Canada to have sailed directly from the country of immigrants’ birth, without stopping along the way (Nayar, 2004). In practice, this primarily affected ships sailing from subcontinental India. The distance between the Indian subcontinent and Canada was such that it was impossible for a ship to sail directly to Canada without stopping to refuel along the way. This in turn made it impossible for ships coming directly from India to dock in Canada. Thus even though Indians were subjects of the British Crown, they were effectively forbidden from traveling to another British colony. The introduction of the Continuous Journey clause subsequent to the Chinatown riots links the histories of Chinese and Punjabi Canadians.

In 1914, Gurdit Singh, a wealthy Punjabi merchant living in Singapore, attempted to help his fellow Punjabis reach Canada by chartering a boat that was to leave from Hong Kong. Three-hundred sixty-four Punjabis arrived to Hong Kong from India by land, and boarded the Japanese steamship called the Komagata Maru bound for British Columbia. However, when they arrived to Vancouver, the Komagata Maru was prevented from docking in Burrard Inlet and they were not permitted to enter Canada. Thus these Punjabi British subjects were forced to return to Asia, despite the fact that they had not expressly violated the Continuous Journey Clause.

Just five years later, another important event happened in Punjabi history. This was the Amritsar Massacre of 1919, which occurred during British colonial rule of the Indian subcontinent in the city of Amritsar of present day Punjab, India (Nayar, 2012). British militia fired on unarmed Punjabis attending an Indian national rally, including children. However, despite Punjabis’ potential desire to leave India at this time, it was difficult for them to travel to Canada in the aftermath of the Komagata Maru incident.

One the other hand, Vancouver Chinese-born Canadians had already formed seven percent of the Chinese community by this time (Ng, 1999). Moreover, in 1921, Vancouver General Hospital dropped its discriminatory policy of sending all Chinese patients to the basement, after the numerous complaints of an infirm elderly Chinese woman (Yee, 2006). However, that same year saw the formation of a second Vancouver Asiatic Exclusion League seeking to end Asian immigration to Canada.
Two years later, on July 1, 1923, the Chinese Exclusion Act was introduced (Ng, 1999; Yee, 2006). Under this act, diplomats, children born in Canada, students, and merchants were the only Chinese permitted to immigrate to Canada. As the name suggests, these restrictive rules did not apply to white European immigrants. Because of this, many Chinese Canadians still refer to July 1 as ‘Humiliation Day,’ and refuse to take part of in Dominion Day celebrations commemorating the formation of a unified Canada.

Furthermore, it required all people of Chinese descent to register with the Immigration Department, regardless of whether they had been born in Canada. The Chinese population of Vancouver did not appear to decline immediately, and had even increased by 1931 (Ng, 1999). However, this was probably due to migration to Vancouver from other parts of BC, and by the 1940s, the Vancouver Chinese population was only a little more than half of what it had been in 1931.

Also during the 1920s, Nayar (2004) reports that informal racism against Punjabis predominated in addition to institutionalized racism. Like the Chinese, Punjabis were restricted from voting, juries, school boards, and the military. In addition, Punjabis faced discrimination in housing, education, and public facilities and services. They were heavily pressured to assimilate to the dominant Anglo-Canadian culture in order to get jobs; it was common for Punjabis during this era to cut their hair, shave their beards, wear Western clothing, and even change their names.

1930 saw the establishment of the Chinese Varsity Club (CVC) at the University of British Columbia (UBC), the very same university where this research was conducted (Chinese Varsity Club, n.d.). It continues to be cultural symbol and place of refuge for Chinese Canadians in their struggle for identity to this day (Ho, 2012). However, Ng (1999) notes that while the CVC was created for Chinese who had been excluded from other UBC activities, it did not become very influential until the 1950s, simply because very few Chinese students even attended UBC until then. For example, only eleven Chinese had graduated from UBC by the mid-1930s.

Chinese Canadians appeared to have strong connections to China during the 1930s, accounting for thirty-two thousand trips to China from 1926 to 1935 (Yee, 2006). As well, they sent over one million Canadian dollars to China between 1937 and 1945 to help them fight the Japanese invasion during the Second Sino-Japanese War, or World
World War II. World War II was good publicity for Chinese in Canada during the 1940s, since China and Canada were allies.

1947 was a historic year for both Punjabi and Chinese Canadians, when Asian Canadians were awarded the right to vote in BC (Yee, 2006; Nayar, 2012). At the same time, the Partition between India and Pakistan occurred, splitting Punjab down the middle and leading to tragedy and violence in the region. Coupled with the tragedies abroad, the electoral enfranchisement of Asian Canadians triggered increased emigration from Punjab to Vancouver during the 1950s.

Around the same time in China, the 1949 Cultural Revolution was stirring, which resulted in the Mao Zedong’s Chinese Communist Party (CCP) defeating the Western-democracy-style Kuomintang, and the latter’s subsequent flight and isolation to Taiwan (Ng, 1999; Yee, 2006). Back in Canada, Chinese Canadians and immigrants faced discrimination by white Canadians due to the Red Scare. As well, Canada officially recognized Taiwan, the Republic of China, as the official Chinese nation, and not the PRC.

Despite McCarthyian xenophobia, the Chinese population of Vancouver doubled between 1951 and 1961, in part because the Cultural Revolution prevented many Chinese in Canada who had wanted to return to China from doing so (Yee, 2006). At the same time, established Chinese Canadians were making an effort to distance themselves from new Chinese immigrants and create their own unique identity (Ng, 1999; Yee, 2006). For example, the CVC began to increasingly be seen as representing exclusively CBCs as opposed to the broader Chinese student body at UBC, though no other Chinese club existed at the time (Ng, 1999).

By the 1960s, the number of foreign and first generation immigrant Chinese students, mostly from Hong Kong, had grown to the point where they were able to establish their own club, the Chinese Overseas Student Association, at UBC (Ng, 1999). From the very start, tension appeared between the two Chinese clubs. For example, despite the cultural trauma of events such as the Nanjing massacre as inflicted by the Japanese during World War II and the shared cultural roots of CBCs and new arrivals from China, the CVC still preferred to co-sponsor events with the Japanese Nisei Varsity Club, an organization of Canadian-born Japanese. This could have been because most of the CBCs in the CVC
during the 1960s would have been born in Canada before the World War II Japanese invasion of China, and thus they may not have shared the resentment towards the Japanese that many of those born in China continue to feel today. Furthermore, the Nisei Varsity Club members were locally born too, and Japanese Canadians had suffered discrimination by the Canadian government as well, particularly during World War II when they were forcibly relocated to internment camps.

The 1960s and 1970s brought in further sweeps of immigration to Vancouver from Guangdong and Punjab, as well as the emergence of Chinese and Punjabi ethnic enclaves in Metro Vancouver (Nayar, 2012). Nayar (2012) qualifies her use of the phrase *ethnic enclaves* in this context by describing those in Metro Vancouver, some of which still exist today, as encouraging enclave members’ physical segregation from the mainstream.

The 1960s also characterize the rise of the Canadian-born Chinese in Vancouver, which was again accompanied by tension between them and incoming first generation immigrants as a result of the Great Leap Forward during 1958-1961 in the PRC (Ng, 1999). In addition, the 1967 amendment to the Canadian Immigration Act enabled Asian applicants to be assessed using the same point system that had previously been employed only for European immigrants. However, in contrast to previous periods, the bulk of Chinese immigration during the 1960s came from Hong Kong and not Guangdong.

The 1970s marked a shift in official ideology, with the government of Canada officially adopted the Multiculturalism Policy in 1971 (Nayar, 2010; Riento, 2013; Yee, 2006). Yet scholars differ in opinion as to how this policy and its later modifications affected immigrant groups in Canada. On the one hand, this policy was rooted in the debate between Francophones and Anglophones in Canada, and its later 1988 version did nothing to directly support languages besides English and French nationwide, as well as Inuktituk in the territory of Nunavut. Riento (2013) argues that this later version shifted any previous focus on integration to one of assimilation. On the other hand, the Multicultural Act did appear to attempt to legitimize the presence of immigrant groups and their descendants in Canada, thereby symbolizing Canada as a more inclusive place. Perhaps this is reflected by the fact that the 1970s were the era of the largest influx of Punjabi immigration (Nayar, 2004).
Furthermore, in 1970, Canada officially recognized the PRC, switching diplomatic recognition of the Chinese nation-state from Taipei to Beijing (Ng, 1999). This, along with the Multicultural Act changed the perception of China in Canada, and fostered renewed association of ethnically Chinese Canadians with Chinese culture. In particular, it gave them the opportunity to re-establish connections with their ancestral region of Guangdong. These ties had been gradually disintegrating up until that point, due to decades of political instability in China, and subsequently severed due to the tense diplomatic relations between China and Canada after the Chinese Communist Party’s defeat of the Kuomintang in 1949.

In 1974, an official PRC consulate was opened in Vancouver, and four years later, the PRC loosened its exit regulations (Yee, 2006). This resulted in the influx of emigration from the PRC to Canada more than tripling between 1978 and 1979. Immigration from Hong Kong also continued to grow during the 1970s, particularly after 1972, when the PRC signaled its intent to regain Hong Kong. In fact Chinese emigration from Hong Kong constituted the largest proportion of Chinese in Vancouver during this period (Ng, 1999; Sandercock & Attili, 2009). As Chinese immigration continued to rise, the tension between newcomers and established Chinese Canadians endured. For example, newcomers criticized CBCs for not speaking Chinese (Ng, 1999).

By the 1980s and 1990s, Punjabi and Chinese ethnic enclaves had been fully consolidated, especially suburban ones (Nayar, 2012; Yee, 2006). Middle-class CBCs were starting to leave Chinatown and move to Richmond and Burnaby, and incoming immigrants were also headed there. Punjabis were located primarily in the cities of Vancouver and Burnaby, though now the highest concentrations in Metro Vancouver are in Surrey. As well, in 1981, over half of all first generation Chinese immigrants had arrived to Canada during the 1970s, and CBCs constituted around a quarter of the Chinese population in Vancouver (Ng, 1999).

In 1984, the Indian government stormed the Golden Temple complex in Amritsar in response to armed separatist activity, again shooting innocent civilians, in what is dubbed ‘Operation Blue Star’ (Nayar, 2004). After this incident, many Punjabi Canadians felt threatened and returned to a traditional Punjabi lifestyle and Sikh religion in order to preserve their cultural heritage. One year later, in 1985, the Canadian government
adopted the Canadian Charter of Rights and Freedoms. Paradoxically, however, discrimination against Punjabis increased that year as a result of the 1985 Air India bombings by Punjabi Sikh separatists, despite the fact that not all Punjabis were in favor of a separate Khalistan or Punjabistan, or the means of violence to establish an autonomous nation.

In 1986, the Canada Employment Equity Act was introduced to combat workplace discrimination (Nayar, 2004). This was also the same year that the Vancouver Khalsa School was founded, which is a private school offering Sikh religion and Punjabi language classes.

The origins of Chinese immigrants shifted in the 1990s from Hong Kong to Mainland China (Yee, 2006). Taken together with the PRC’s continual promotion of Putonghua, or Standard Mandarin, in addition to the fact that many Chinese immigrants were coming from outside of Guangdong, this meant the rise of Mandarin over Cantonese in Vancouver, a pattern that persists today. The finalization of the Sino-British agreement to return Hong Kong to China occurred in 1997, but by this point fears of a total CCP takeover had ebbed in Hong Kong.

The 1990s also was filled with Canadian and British Columbian recognition of Punjabis as an integral part of Canadian history, and multiple actions were taken to demonstrate this. In 1990, the RCMP officially permitted officers to wear turbans, a common marker of Sikh identity (Nayar, 2004). In addition, the Canadian government issued an apology for the Komagatu Maru incident, and a plaque was placed at Portal Park in Vancouver. 1993 inaugurated the official recognition of the city of Vancouver’s Punjabi Market, with bilingual English-Punjabi signs being put up in the area (Nayar, 2004). Two years later, in 1995, British Columbia officially recognized the Vaisakhi parade commemorating the Punjabi harvest festival.

Today, the predominantly Cantonese Chinese and Punjabi communities are still alive and well. In particular the Strawberry Hill and Newton areas in Surrey and Delta, respectively, remain Punjabi ethnic enclaves, and Richmond remains a Chinese ethnic enclave (Nayar, 2012). Additionally, the Punjabi Market, also known as ‘Little India,’ in Vancouver’s Sunset district, and Chinatown, east of Vancouver’s downtown, constitute less insular communities as well as symbols of Cantonese and Punjabi Canadian heritage.
And while clearly the number of Mandarin-speakers is catching up with the number of Cantonese-speakers, Cantonese is still clearly alive and well, seeing as that it is the second-most common non-official home language and mother tongue in Metro Vancouver after Punjabi.

However, Punjabis and Cantonese Canadians still struggle for identity. This is reflected in Nayar’s (2004) work suggesting that many Punjabi Canadians do not appreciate the way they have and continue to be portrayed in the mainstream media. In addition, there is tension between Punjabi and Chinese immigrants who arrived to Canada in adulthood or late adolescence and Punjabis and Chinese raised in Canada. For example, younger Vancouver-born Punjabis often clash with their Punjabi-born parents and grandparents by renegotiating the value of traditional cultural notions such as strong adherence to izzat, i.e. preserving honour and evading shame. Nayar (2004) reports that while the older Punjab-born generations continue to use strategies to save face as avoiding open discussion about oneself and one’s family, the younger Canadian-born Punjabis tend to view izzat more negatively and as a hindrance to the resolution of community problems.

The video released in 2008 by the Chinese Varsity Club portrays the tension between already established Chinese Canadians and more recent immigrants from Hong Kong (Ho, 2012). The video contrasts depictions of CBCs as trendy and fashionable with depictions of so-called ‘Hongers’ as nerdy and not proficient in English. Predictably, the video raised an outrage. The club responded by apologizing for the video and was forced by the UBC student government to undergo equity training.

1.3 Research problem

Inspired by previous studies and the rich history of Punjabi-Canadians and Chinese-Canadians in Vancouver, the present research sought to investigate whether Vancouverites from East Indian, Chinese, and European ethnolinguistic backgrounds would exhibit differences in their English speech. However, there exists a paucity of ethnolinguistic research in Vancouver, and findings from other geographical regions
support the view that multiple linguistic differences may be exhibited in the speech of
different ethnic groups (e.g. Clopper & Bradlow, 2009). Thus prior to investigating
potential differences in specific linguistic variables between groups, it was decided to
explore whether Vancouver listeners are able to perceive differences between the groups.
One can reason is that if listeners perceive the speech of these three groups differently,
differences in speech production must exist as well, which the listeners must be using as
cues in order to distinguish each group.
2 Methodology

The basic design of the study was as follows: record talkers, present sentences from their recordings to listeners, and have listeners identify talker ethnicity for each sentence in a forced-choice identification task. In the production study, or ‘talking,’ subjects were recorded narrating a video depicting a story they had just read. In the perception study, or ‘listening,’ different subjects were presented with sentences from the talkers’ video narrations and asked to identify the ethnicity of the talker saying each sentence. All participants received verbal instructions by the same researcher, and all participants completed their part in the study by filling out a background questionnaire.

2.1 Talking

2.1.1 Talkers

Talker eligibility criteria were as follows:

(1) Talkers needed to be native speakers of English or to have acquired English by the age of six years.
(2) Talkers must have received all education since kindergarten in English.
(3) Talkers must not have had any speech, language or hearing disorders.
(4) Talkers must have been between the ages of eighteen and twenty-five years at the time they participated in the study.
(5) Talkers must have been raised in Metro Vancouver since age eight.
(6) Talkers must have ethnically self-identified as Chinese, white, or East Indian.
(7) Talkers who ethnically self-identified as Chinese or East Indian must have reported the native language of their ancestors to be Cantonese or Punjabi, respectively.
(8) Talkers who ethnically self-identified as white must not have reported a home language other than English.
Several strategies were employed to ensure that participants met eligibility criteria. Firstly, it was confirmed that potential participants met eligibility requirements before taking part in the study by asking them questions regarding each requirement (e.g., “Where were you raised from ages eight onwards?”). Secondly, researchers took note of any indications of speech, language, or hearing disorders in talkers’ recordings. Thirdly, participants’ background questionnaires were reviewed after each participant’s completion of the study.

Talkers were recruited using the following strategies: website and mailing list advertising through the University of British Columbia Psychology Graduate Council Paid Participants Studies, poster advertisements, class announcements, and the snowball method. Most talkers were recruited from the student population of the University of British Columbia (UBC) in Vancouver. However, due to unknown recruitment difficulties, most of the male Punjabi talkers were recruited from Simon Fraser University in Surrey. The study was advertised as having the title “Reading aloud and listening” in order to avoid revealing the true purpose of the study and thus minimize potential top-down effects that might have influenced talkers’ speech if they had known that the study concerned ethnicity.

Thirty-seven talkers in total were recorded. Of these thirty were included in the study, and seven were excluded, based on eligibility criteria. The first talker was excluded because they were part of the pilot study, and modifications were made to the instructions after they completed the study. Another talker was excluded because their recordings provided auditory evidence that they might have had a speech disorder. Two talkers were excluded, because they did not finish watching and narrating the video. One talker was excluded because they were Chinese and identified Mandarin, but not Cantonese, as the language of their ancestors. Another talker was excluded, because their background questionnaire demonstrated that they had not been raised in Metro Vancouver since the age of eight. Finally, another talker was excluded because their background questionnaire showed that they were over the age of twenty-five.

Ten talkers from each ethnic group comprised the thirty talkers included in the study. There were equal numbers of males and females from each ethnic group. All talkers met eligibility criteria except two, due to researcher oversight. One of these was a male East
Indian talker who was raised in Terrace, British Columbia, which is not part of Metro Vancouver. The other was a white female talker whose home language was French. These two talkers met all other eligibility criteria.

2.1.2 Procedure

Prior to collecting data, participants were given a consent form and asked to sign it if they agreed. At this point in the procedure, talkers were not told the purpose of the study. They were only told that they would be recorded reading a story and watching a video while narrating it.

After receiving verbal instructions, they were then recorded reading the story “The North Wind and the Sun” aloud (read speech) and subsequently narrating a video depicting the same story (semi-spontaneous speech). A copy of “The North Wind and the Sun” can be seen in Appendix A. Sentences from their semi-spontaneous speech served as the stimuli for the perception study. The purpose of the read speech was to familiarize the subjects with the story in the hopes of minimizing disfluencies common in spontaneous speech (Thomas and Carter, 2006).

Recording was done using the software Audacity and AKG C520 condenser microphone headsets. Typed instructions, the story, and the video were presented using the psychology study softwares E-Prime (Schneider, Eschman, & Zuccolotto, 2002a; Schneider, Eschman, & Zuccolotto, 2002b) and PsychoPy (Peirce, 2007). Microphone position and gain level of the recorder were adjusted appropriately for each talker. One talker requested not to wear the microphone on his head, so the talker was permitted to hold it manually.

After the recording, talkers were instructed to complete a handwritten background questionnaire consisting of questions regarding the language backgrounds of them and their families; proficiency level in languages they spoke; socioeconomic and educational backgrounds; and residential history of them and their families. A full version of the questionnaire can be seen in the Appendix B. After completing the background questionnaire and receiving ten Canadian dollars in compensation, talkers were debriefed.
2.1.2.1 Collection of read speech

Talkers were instructed to read the story “The North Wind and the Sun” silently to themselves in order to familiarize themselves with it. Next they were instructed to read the story out loud. They were permitted to read the story to themselves as many times as they liked, but they were only permitted be recorded reading the story aloud once. This story was chosen because it has been used widely in linguistic research and is not directly related to race, language, or ethnicity (e.g. Grabe & Low, 2002; Mok & Dellwo, 2008).

2.1.2.2 Collection of semi-spontaneous speech

Video narrations were chosen for the semi-spontaneous in order to provide a compromise between ecological validity and control. Using completely spontaneous conversational speech might introduce confounding factors such as interlocutor or topic (Meyerhoff, 2006). On the other hand, using more formal read speech or word lists might not be representative of the way talkers used language in everyday life; for example, read speech and word lists constrain the grammatical and lexical variants that talkers might employ. In addition, providing a video for talkers to narrate helps to entertain talkers while hopefully lightening the memory load of the task.

Talkers were instructed to narrate a video from the National Film Board of Canada, which depicted the same story, “The North Wind in the Sun,” that they had just read (Leyer, Drew, & Verrall, 1972). They were instructed to narrate the video as they watched it and use their own words in their narrations. The video was presented to them with the sound off. They were told that if they forgot the story, they could just describe what was happening in the video instead. They were only permitted to narrate the video once, unless technical issues arose which froze the video. In the event of the video freezing, subjects were permitted to re-narrate the video, and it was attempted to include sentences from their first narrations only for the perception study.
2.2 Listening

2.2.1 Listeners

The listener eligibility criteria were as follows:

(1) Listeners needed to be native speakers of English or to have acquired English by the age of six years.
(2) Listeners must have received all education since kindergarten in English.
(3) Listeners must not have had any speech, language or hearing disorders.
(4) Listeners must have been between the ages of eighteen and twenty-five years at the time they participated in the study.
(5) Listeners must have been raised in Metro Vancouver since age eight.
(6) Listeners must not have had participated in the production study.

The same strategies employed in the production study to ensure participants met eligibility criteria were utilized in the perception study.

Listeners were recruited using the snowball method and website and mailing list advertising through the University of British Columbia Psychology Graduate Council Paid Participants Studies. The study was advertised with the title “Reading aloud and listening II.”

Sixteen listeners participated in the perception study. All listeners met eligibility criteria.

2.2.2 Stimuli

The minimum number of sentences produced by a single talker was six. For this reason, each listener was presented with six sentences from each talker. Because listeners were exposed to stimuli from all thirty talkers, each listener was presented with one hundred eighty sentences in total, excluding the practice ones, which are described below.

Sentences from talkers’ video narrations were screened using certain criteria in order to be considered for use in the perception experiment. The criteria are ranked in order of their importance, with criterion (1) being the most important. They are as follows:
(1) If a talker had to re-watch part of the video due to technical issues, sentences that narrated the re-watch part were only included from the first viewing.

(2) Sentences needed to minimally consist of a main verb and a subject noun.

(3) Sentences needed to be less than twenty seconds long in total duration.

(4) Sentences must not have contained pauses over two seconds long.

Because some talkers did not produce enough sentences meeting all criteria, some of the lower ranked (less important, e.g. criterion four) criteria were violated. Four sentences were included with pauses over two seconds, each one being from a different talker.

In order to account for the fact that many talkers produced more than six sentences, two sets of stimuli were created, with each set being presented to half of the listeners. For each set of stimuli, six sentences were selected per talker. All six sentences were selected for each talker that only produced six sentences, or only produced six sentences meeting criteria. For each talker that produced at least twelve sentences meeting criteria, six sentences were randomly selected per talker for each set of stimuli, with completely different sentences used between sets. For talkers producing more than six but less than twelve sentences meeting criteria, six sentences were randomly selected for the first set of stimuli, and the second set included all sentences not included for the first set, with the remainder of sentences being a random selection of those already included in the first. In this way, more than six sentences could be included per talker in the perception study, but each listener was only exposed to six sentences per talker, and the difference in the number of listener judgments per talker would be minimized despite different talkers producing more or less sentences meeting criteria.

In addition, six sentences provided stimuli for the practice round. A white twenty-four-year-old female talker spoke three sentences, and twenty-six-year-old Punjabi male talker, who had been originally excluded from the study due to the age criterion spoke another three. Except for the fact that the male talker was twenty-six, both practice talkers would have met all talker eligibility criteria, including being raised in Metro Vancouver. None of the sentences from the practice round were used in the analysis.

All stimuli, including practice stimuli, were normalized for peak amplitude.
2.2.3 Procedure

As in the production study, listeners were given a consent form and asked to sign it if they agreed.

Then the participants were told that they would be presented with sentences and would have to identify the talker’s ethnicity for each sentence as quickly and accurately as possible. They were instructed to wait until the sentence had finished playing to make their judgments, and visual prompts were used to ensure that listeners would know when this occurred. Each sentence was only presented to the listener once. They were also instructed to rate their confidence in their identification choice for each sentence, on a scale of one to five, with one being “just guessing” and five being “100% confident.”

Listeners completed a practice round before their responses were logged. The practice round consisted of the six practice sentences specified in the previous subsection of this paper. The practice round was provided to ensure that no technical issues arose, as well as to ensure that the presentation intensity level was neither inaudible nor painful. No participants reported that the presentation intensity level was inaudible or painful; thus the intensity level was constant across listeners.

All sentences, including the practice, were presented to them using the software E-Prime (Schneider, Eschman, & Zuccolotto, 2002a; Schneider, Eschman, & Zuccolotto, 2002b). E-Prime was also used to provide written instructions and log listener responses. Listeners were given five seconds to make each confidence and identification judgment, and made these judgments using a button box. The five buttons on the button box were labeled one through five. For the confidence judgments, listeners were instructed to press the button that was numbered in accordance to the confidence judgment they wanted to make. For example, if a listener felt “100% confident” in their identification judgment, they were instructed to select the button labeled with the number five. For the identification judgments, each of the ethnicity choices was assigned to a button numbered either one, three, or five, but the specific ordering of the choices was counterbalanced across listeners such that there were six different possible button assignments. For example, for some participants, button one corresponded to “East Indian,” button three corresponded to “Chinese,” and button five corresponded to “white;” for other
participants, button one corresponded to “white,” button three to “East Indian,” and button five to “Chinese.” Listeners could only make their judgments once for each sentence.

After listeners completed all one hundred eighty trials of the perception study, they were instructed to complete a handwritten background questionnaire. The full questionnaire can be seen in Appendix C. Questionnaire content was the same for both the perception and production study, except that the listener questionnaire contained one additional item. This additional item asked listeners to rank Chinese, East Indian, and white ethnic groups according to their frequency of interaction with that group. The ranking scale was from one to three, with one corresponding to the group they interacted with most, and three corresponding to the group they interacted with least. Listeners were permitted to give groups equal ranks if they interacted with them equally often. From hereon, I use the term (listener) network rank to denote the ethnic group which a given listener ranked as number one.
3 Results

Listener data was analyzed using the software R (R Core Team, 2014). Mean proportion correct, or mean accuracy rate, across all sixteen listeners was calculated by finding each listener’s mean accuracy rate, then finding the mean of all listeners’ mean accuracy rates. In other words, a grand mean was used incorporating within-listener means. In all graphs below, error bars represent the mean plus or minus one standard error, with standard error being equal to the square root of the variance of the mean accuracy rate divided by the square root of the total number of sentences for which listeners identified talker ethnicity.

Table 1, below, shows values for mean accuracy rates. Since none of the sixteen listeners identified their listener network rank as East Indian, only accuracy rates for Chinese and white listener network ranks are shown. Seven listeners reported Chinese network ranks, and nine reported white network ranks.

Table 1: Listeners’ mean talker ethnicity identification accuracy rates, and their corresponding standard errors.

<table>
<thead>
<tr>
<th>Listener Network Rank</th>
<th>Talker ethnicity</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chinese</td>
<td>East Indian</td>
</tr>
<tr>
<td>Chinese</td>
<td>0.688±0.047</td>
<td>0.469±0.036</td>
</tr>
<tr>
<td>White</td>
<td>0.531±0.033</td>
<td>0.498±0.042</td>
</tr>
<tr>
<td>Overall</td>
<td>0.600±0.033</td>
<td>0.485±0.028</td>
</tr>
</tbody>
</table>

As seen in Figure 2, below, the results suggest that listeners were able to identify talker ethnicity far above chance levels. Due to the large magnitude of the difference between listener accuracy rates and chance accuracy rates, hypothesis testing was not deemed necessary to test the null hypothesis that listener accuracy rates were at chance levels against the alternative hypothesis that listener accuracy rates were not at chance levels.
Figure 2: Listeners’ mean talker ethnicity identification accuracy rates (mean proportion correct) by talker ethnicity. Error bars represent plus or minus one standard error, and the dashed line denotes chance performance, i.e. proportion correct=1/3.

To investigate the relationships between listener identification accuracy, talker ethnicity, and listener network rank, mixed effects binomial logistic regression models was fitted to the data using the command `glmer` in the R package `lme4` (Bates, Maechler, Bolker, & Walker, 2014). Fixed effects were talker ethnicity and listener network rank. The interaction between talker ethnicity and listener network rank was used as a by-subject random slopes.
Treatment-style contrasts were used to compare performances between different listener network rank groups when identifying different talker ethnicity groups. This means that for each factor, there is a reference level to which the other levels are each compared with respect to the dependent variable, and the levels that are not reference levels are not compared to each other. Contrasts correspond to all fixed effects that are not the intercept, since non-intercept fixed effects denote comparisons being made. The intercept coefficient corresponds to accuracy for listeners with the given reference level network rank when identifying talkers of the given reference level ethnicity.

Due to the fact that factor levels that are not reference levels are not compared to each other, as well as the possibility that contrast significance may depend on the reference level used, six models were fit, one per each combination of unique talker ethnicity and listener network rank reference levels. Contrasts are considered significant here only if they were found to be significant in all models that tested them. In this section, results are only reported for the three models, because these models together can sufficiently show which contrasts are significant. These results can be seen in Tables 2-4, below.

**Table 2:** Model 1 for regressing listeners’ accuracy identifying talker ethnicity against the fixed effects of talker ethnicity (“ethnicity”) and listener network rank (“rank”). Random slopes were used for talker ethnicity and network rank as well, but they are not shown here. Talker ethnicity=Chinese was used as the reference level for talker ethnicity, which is the intercept, and network rank=Chinese was used as the reference level for listener network rank. A single asterisk indicates significance at \( p < 0.05 \) only in this model. Two asterisks indicate significance at \( p < 0.05 \) across all six models.

| Fixed effect             | Coefficient estimate | Std. Error | z value | \( p(>|z|) \)   |
|--------------------------|----------------------|------------|---------|----------------|
| intercept                | 0.8218               | 0.1976     | 4.158   | 3.20E-05*      |
| ethnicity=East Indian     | -0.9479              | 0.2133     | -4.444  | 8.82E-06**     |
| ethnicity=White           | -0.0299              | 0.3188     | -0.094  | 0.92524        |
| rank=White               | -0.6928              | 0.2344     | -2.955  | 0.00312**      |
| ethnicity=East Indian:rank=White | 0.8072              | 0.3020     | 2.673   | 0.00752*       |
| ethnicity=White:rank=White| 1.0290               | 0.3768     | 2.731   | 0.00632*       |
Table 3: Model 2 for regressing listeners’ accuracy identifying talker ethnicity against the fixed effects of talker ethnicity (“ethnicity”) and listener network rank (“rank”). Random slopes were used for talker ethnicity and network rank as well, but they are not shown here. Talker ethnicity=White was used as the reference level for talker ethnicity, which is the intercept, and network rank=Chinese was used as the reference level for listener network rank. A single asterisk indicates significance at \( p<0.05 \) only in this model. Two asterisks indicate significance at \( p<0.05 \) across all six models.

| Fixed effect                      | Coefficient estimate | Std. Error | z value | \( p(>|z|) \) |
|-----------------------------------|----------------------|------------|---------|----------------|
| intercept                         | 0.7919               | 0.1961     | 4.039   | 5.38E-05*      |
| ethnicity=East Indian              | -0.9180              | 0.2565     | -3.579  | 0.000345**     |
| ethnicity=Chinese                  | 0.0300               | 0.3187     | 0.094   | 0.925037       |
| rank=White                        | 0.3362               | 0.2590     | 1.298   | 0.194316       |
| ethnicity=Punjabi: rank=White     | -0.2218              | 0.3697     | -0.600  | 0.548491       |
| ethnicity=Chinese: rank=White      | -1.0290              | 0.3767     | -2.732  | 0.0063**       |

Table 4: Model 3 for regressing listeners’ accuracy identifying talker ethnicity against the fixed effects of talker ethnicity (“ethnicity”) and listener network rank (“rank”). Random slopes were used for talker ethnicity and network rank as well, but they are not shown here. Talker ethnicity=White was used as the reference level for talker ethnicity, which is the intercept, and network rank=White was used as the reference level for listener network rank. A single asterisk indicates significance at \( p<0.05 \) only in this model. Two asterisks indicate significance at \( p<0.05 \) across all six models.

| Fixed effect                      | Coefficient estimate | Std. Error | z value | \( p(>|z|) \) |
|-----------------------------------|----------------------|------------|---------|----------------|
| intercept                         | 1.1281               | 0.1693     | 6.662   | 2.69E-11*      |
| ethnicity=Punjabi                 | -1.1398              | 0.2664     | -4.279  | 1.88E-05**     |
| ethnicity=Chinese                 | -0.999               | 0.2009     | -4.972  | 6.62E-07**     |
| rank=Chinese                      | -0.3362              | 0.2591     | -1.297  | 0.19456        |
| ethnicity=Punjabi: rank=Chinese   | 0.2218               | 0.3699     | 0.600   | 0.5488         |
| ethnicity=Chinese: rank=Chinese   | 1.0290               | 0.3769     | 2.730   | 0.00633**      |

Five contrasts reach significance across all six models. Firstly, Chinese-ranked listeners were significantly more likely to accurately identify Chinese talkers than East Indian talkers. Secondly, Chinese-ranked listeners were significantly more likely to accurately identify white talkers than East Indian talkers. Thirdly, Chinese-ranked listeners were significantly more likely to accurately identify Chinese talkers than white-ranked listeners were. Fourthly, white-ranked listeners were significantly more likely to accurately identify white talkers than Chinese talkers. Lastly, white-ranked listeners were
significantly more likely to accurately identify white talkers than East Indian talkers. These results can be visualized in Figure 3, below, which demonstrates the effect that the interaction between listener network rank and talker ethnicity has on identification accuracy.

Figure 3: Listener identification accuracy by listener network rank. The horizontal dashed line indicates chance performance, and error bars represent plus or minus one standard error.

Table 5, below, is a confusion matrix for listener responses against actual talker ethnicities. Figure 4, below, visualizes this confusion matrix; cases in which listeners who did not respond are not included, as these constituted less than 1.5% of all responses overall and for each listener group.
Table 5: Confusion matrix data for listeners’ talker ethnicity identification responses. ‘Talker ethnicity’ indicates the correct response, or actual talker ethnicity, whereas ‘Listener response’ indicates listener judgments of talker ethnicity. Un-parenthesized values are counts; parenthesized values are approximate relative frequencies in percentage format.

<table>
<thead>
<tr>
<th>Listener response</th>
<th>Talker ethnicity</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chinese</td>
<td>White</td>
</tr>
<tr>
<td>Chinese</td>
<td>577 (20%)</td>
<td>121 (4.2%)</td>
</tr>
<tr>
<td>White</td>
<td>283 (9.8%)</td>
<td>692 (24.0%)</td>
</tr>
<tr>
<td>East Indian</td>
<td>89 (3.1%)</td>
<td>132 (4.6%)</td>
</tr>
<tr>
<td>No response</td>
<td>11 (0.4%)</td>
<td>14 (0.5%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>960 (33.3%)</strong></td>
<td><strong>960 (33.4%)</strong></td>
</tr>
</tbody>
</table>

Figure 4: Confusion matrix for listeners’ talker ethnicity identification responses. The x-axis represents actual talker ethnicity; the y-axis represents listener’s judgments of talker ethnicity. Darker colors represent lower response frequencies, and lighter colors represent higher response frequencies.
Listeners were most likely to identify talkers as white and least likely to identify talkers as East Indian, regardless of actual talker ethnicity. As well, the confusion matrix implies that listeners most often chose the correct talker ethnicity, as suggested earlier by Figure 2 and the regression model. In cases where listeners were inaccurate, they were more likely to misidentify both East Indian and Chinese talkers as white than as the other visible minority ethnicity, although this difference appears larger for Chinese than East Indian talkers. It does not appear that listeners were more likely to misidentify white talkers as Chinese or East Indian.
4 Discussion

4.1 Findings

In general, results suggest that listeners are able to identify talker ethnicity above chance levels overall. However, listeners with different network ranks did not appear to be equally good at identifying each talker ethnicity. While both groups appeared to be better at identifying white talkers than East Indian talkers, patterns at identifying Chinese talkers seemed to diverge between listeners with different network ranks. White-ranked listeners appeared better at identifying white talkers than Chinese talkers, and exhibited little difference in accuracy identification between Chinese and East Indian talker groups. On the other hand, Chinese-ranked listeners seemed to identify white and Chinese talkers equally well. Finally, talkers appeared more likely to misidentify East Indian and Chinese talkers as white rather than as the other visible minority group, and white talkers were equally misidentified as either East Indian or Chinese.

The above-chance level of talker ethnicity identification across different talker and listener groups suggests that there must be differences in the features speakers of different ethnicities use, which listeners use in making their judgments. This provides evidence for the existence of Cantonese Chinese and Punjabi East Indian ethnolinguistic repertoires in Vancouver English. I argue that this is an especially appropriate inference to make considering that we controlled for the confounding effects of interlocutor and topic. That is, perhaps we can consider the recordings here as sort of a ‘baseline’ for the talkers. In addition, this approach is probably a more robust as well as time- and cost-effective alternative to attempting to rule out potential variables before doing any perceptual experiments. As well, it is line with previous theories on language contact with respect the characteristics of the Metro Vancouver environment setting up several optimal conditions for the formation of Punjabi East Indian and Cantonese Chinese ethnolinguistic repertoires.

Additionally, the hypothesis that experience aids talker ethnicity identification explains the discrepancies between listener groups when identifying Chinese talkers.
More specifically, Chinese-ranked listeners probably did better identifying Chinese talkers than the white-ranked listeners, because they were more exposed to their speech. The advantage of familiarity also may explain the lower accuracy rates for East Indian talkers, considering that no listeners had East Indian network ranks. That is, perhaps listeners exhibited poorer performance at identifying East Indian talkers because they had less experience with them, and therefore they had less expectations of what an East Indian talker sounds like.

On the other hand, it is curious that listeners from both network ranks appeared to have equally accurate identifications of white talkers, and that Chinese-ranked listeners were no better at identifying Chinese talkers than white talkers. Clearly network rank cannot explain this pattern. However, experience may still account for this pattern if one considers the dominance that white people have had in Metro Vancouver throughout history and probably continue to have even today. If the white Vancouverite is promoted as the “normative” Vancouverite, presumably white English will not only be most accessible to all Canadians, but familiarity with white English will also most imperative. Thus regardless of one’s background, one will have no choice but to be exposed to white English in Vancouver.

The confusion matrix suggests that listeners misidentify talkers from both minority ethnicities as white more often than as the other minority ethnicity. There are at least two potential implications of this finding, although it is important to note that significance tests are yet to be conducted. Firstly, the results fail to provide evidence for an overall ‘Asian’ or ‘ethnic minority’ variety of English in Vancouver. This is opposed to other studies that have found evidence of English repertoires shared between various minority ethnicities; see Kerswill (2010) for examples.

Secondly, it is possible that our confusion patterns imply listeners tend to assimilate talker ethnicity to that of the dominant ethnic group when they misidentify the ethnicity of talkers from East Indian and Chinese ethnic backgrounds. There could be multiple reasons for this. For one, it could be that speech differences between talkers of different ethnicities were minimal, and that is why listeners get talkers confused. In other words, the features differentiating the visible minority talkers from the white talkers are not that perceptually salient.
I do not consider this a solely production-based explanation likely, however. First of all, if talkers’ productions accounted for talkers’ misidentification patterns alone, one would have expected one of two scenarios diverging from that seen here. One of these entails a situation in which all talker groups are equally similar to each other in production. If this were truly the case, one would not expect to see any clear misidentification patterns. Crucially, this would mean that the rate of misidentifying Chinese talkers as white would be the same as misidentifying Chinese talkers as group East Indian, which is clearly not the case. The other scenario is one in which two talker groups speak similarly to each other but the other group speaks differently. This cannot account for listener’s confusion patterns either, because then we would expect that *either* Punjabi *or* Chinese talkers would be more likely misidentified as white, but not both. One also would have expected that this pattern would be bidirectional, in that white talkers would be more commonly misidentified as the group they were most similar with in terms of production. Finally, if talkers really were very similar to each other, one would have anticipated lower accuracy rates, as listeners would not have had substantial indexical information to extract from the signal.

Therefore explanations for listener response patterns must lie in the perceptual arena. It could be that when listeners are uncertain of a talker’s ethnicity, they just make reference to the dominant norm. We will have to look at the confidence responses and reaction times to assess this claim, which has not yet been done. Another explanation is that, regardless of their certainty, listeners may assimilate talkers to the dominant norm when there are not enough perceptually salient cues for listeners to extract them in the signal. Again, this is a production-based explanation, but it also incorporates perception, as without assimilation patterns, one could not account for the patterns observed here.

Taking the interaction between network rank and into account with misidentification patterns suggests that Cantonese and Punjabi Canadians’ ethnolinguistic repertoires may be less perceptually available to out-group members. Whether unconscious or not, this limits the accessibility of the marked elements of their speech to out-group members and thereby protects them from discrimination. Thereby ethnolinguistic repertoires may provide a mechanism for minorities to preserve their cultural heritage in a ‘safe space.’ However, further analyses are needed to determine
whether misidentification patterns reach significance, as well as whether there is an interaction between listeners’ confusion patterns and their network rank.

4.2 Potential limitations

Despite my clarification and reasoning behind the labels used here, it may still be argued major limitation of this study is the use of ethnic and racial labels. These labels are most likely not universally politically correct, nor uncontroversial accurate in various fields of academia, particularly in ones concerned with race and ethnicity specifically, such as anthropological ethnography. In addition, it is possible that they may have influenced listener identifications in undesired ways.

Firstly the term ‘white’ is extremely controversial and ambiguous. One of the problems with the word ‘white’ is its exclusivity. I will give myself as an example. My mother considers herself ‘white,’ and is of predominantly Scottish and German ancestry, whereas my father considers himself ‘Chinese,’ and his parents are first-generation immigrants from Guangdong. However, my white classmates always considered me as only Chinese or Asian, despite the fact that they identified my mother as white. Therefore it can be argued that ‘white’ is like the idealization of the colour white, denoting some sort of purity that will be tainted by any other colour or race. In addition, as Bucky, a cat strongly resembling my cat from the cartoon Get Fuzzy, claims, ‘white’ people are clearly not white in colour, unless they have some sort of genetic albinism or a lack of melanin attributable to other causes (Conley, 2002). Bucky likes to say that ‘white’ people, such as his owner, Rob, are ‘pink,’ and he frequently teases Rob for being pink. Bucky also describes Rob’s African American friend as ‘mocha.’

It can be said that the other two labels used here are mired in just as much controversy as ‘white.’ Using the term East Indian to refer to Punjabis may imply that Punjabis are only from India, and that ignores the fact that there is another province called Punjab heavily populated by self-identified Punjabis in neighboring Pakistan. Moreover, the distinction between India and Pakistan was actually imposed by the British, and they drew the line based on religion, ignoring the fact that the territory
inhabited by Hindus, Sikhs, and Muslims in the region overlapped. Even worse, the region of Punjab straddled the border, and thus the 1947 Partition was an extremely traumatic event for many Punjabis. The Partition and the subsequent resentment of many Punjabis towards the Indian government for not granting them “special consideration as a minority group” spurred militant separatist movements during the 1980s and 1990s (Guha, 2003, p. 31, as cited by Nayar, 2012). The successive violent reactions of the Indian government resulting in the death and injury of many civilian Punjabis, including children, also did not serve to foster an East Indian association in many Punjabis. In summary, requiring that Punjabis be ‘East Indian’ as an eligibility criterion may have excluded Pakistani Punjabis, and this term is also a sensitive one based on imperialist and national ideology.

The label ‘Chinese’ also suffers from contentious underpinnings as well. First of all, using the term ‘Chinese’ does not delineate whether one is referring to ethnic Han Chinese, considered being the majority of the PRC and Hong Kong in terms of prestige and proportion, as opposed to the numerous other ethnic minorities of China who are Chinese citizens but are not considered Han (Chen, 2010). In addition, while China as a nation state has arguably existed for thousands of years, there is considerable cultural and linguistic diversity even among Sinitic language speakers who consider themselves Han Chinese, which is considered the dominant ethnic group of China. Furthermore, the label Han Chinese is at least as ill defined, controversial, and divergent as the label white, with similar ethnocentric colonially motivated associations (Chen, 2010). Despite this, the notion of a unified Chinese race/ethnic group continues to exist, and Chen (2010) claims that the word ‘Chinese’ is increasingly being substituted for ‘Han’ globally.

Another potential issue with using nation-associated labels for the visual minorities and a broad race-based label for the dominant group is perhaps this implies an association between the former groups and ‘foreignness’ or perhaps even the general ‘other,’ as their labels are based on the names of foreign countries. That is to say, perhaps listeners might be identifying talkers according to how ‘non-native’ or ‘non-standard’ they sound. While it is certainly important that we as researchers should not be perpetuating discriminatory stereotypes by projecting them onto listeners, listeners may use this distinction to make their categorizations regardless of which labels are used, and
perceptions of ethnicity and nativeness in speech may not be inseparable. In addition, if listeners were truly using ‘non-nativeness’ or ‘foreignness’ as a cue, we would have expected listeners to confuse talkers of the visible minority ethnicities with each other in their identifications more than with white listeners, which clearly did not appear to be the case.

As well, it is unclear as to whether better labels exist than the ones used here. The complexity of race and ethnicity as elucidated by the previous paragraphs just serves to point out that choosing which labels to use is a daunting task. In addition, standards of political correctness and consensus among academics are constantly changing and variable between different places in the world, and different fields as well, making it hard to establish criterion determining the appropriateness of different labels. For example, we could have used ‘Indo-Canadian’ instead of ‘East Indian,’ but there is already ethnographic evidence that Punjabis find the former label offensive, and it has been claimed that people from India tend to identify more strongly with their regional language group or religion (Nayar, 2004; Nayar, 2012). We also could have substituted South Asian, which avoids the issues associated with Indian national politics. This may have been a better alternative for confirming eligibility in order to make sure that we included Punjabis from the current territory of Pakistan, and it was indeed employed on the background questionnaire. However, the fact that one of the listeners self-identified as South Asian but reported being a native speaker of Vietnamese makes it clear that South Asian is far too broad of a term to have presented as a category for listeners.

Instead of ‘white,’ we could have used ‘European,’ or ‘Caucasian.’ Yet I do not think that ‘European’ is a commonly used term in self-identification, particularly if one is not born in Europe. While it appears to me that people may say they are of ‘European descent,’ they may not say, “I am European.” I have heard the term European used in a North American English context more commonly to refer to foreigners who are visiting from Europe or have moved to Canada as adults. In addition, I have heard some of my friends who were born and raised in Europe specifically say that they do not consider white Canadians and Americans who were not born and raised in Europe to be European, especially if their European descent goes back for more than a few generations. Additionally, ‘European’ involves some of the same issues as ‘white’ and ‘Caucasian’ do
in that it is not clear who is included. For example, are Turkish and Russian people considered European or are they considered Asian? ‘Caucasian,’ for its part, is in my opinion in even worse and inaccurate term than ‘white.’ Unlike the other ethnic/racial terms mentioned in this section, it is documented that ‘Caucasian’ was actually coined by one person, Christoph Meiners, who in 1785 proposed a taxonomy comprised of only two races, ‘Caucasoid’ and ‘Mongoloid’ (Baum, 2008). In addition, Meiners claimed that Caucasians were both more physically attractive and morally virtuous than those he dubbed ‘Mongolians.’

Alternatives to using Chinese could have included Cantonese, perhaps. However, as mentioned previously, ‘Cantonese’ involves some of the inclusivity problems as ‘Chinese.’ Furthermore, although even speakers of non-standard Yue varieties might refer to themselves as ‘Cantonese’ speakers, I have never observed anyone referring to either themselves or other people as ‘Cantonese’ or ‘Yue’ with respect to ethnicity. That is to say, speakers of Yue language varieties may say, “I speak Cantonese,” but I have never heard anyone say, “I am Cantonese.”

The latter is true for Yue as well. I have never even heard anyone say they that they spoke ‘Yue,’ and my father did not even know that Toisan was considered a Yue language until I told him. If native speakers do not know what ‘Yue’ means, how can we expect listeners to know? It is also unclear whether listeners who are not South Asian or Chinese would be familiar with the terms ‘Cantonese’ and ‘Punjabi.’ Or even if those listeners were familiar with those terms, it is possible that they might not be familiar enough for them to have salient representations of them associated with ethnolinguistic repertoires.

Furthermore, numerous studies have found that naïve listeners tend to perform poorly at talker categorization tasks when the categories are overly specific in nature (Clopper & Pisoni, 2005), and that they do not appear to have cognitive maps corresponding to linguists’ maps of repertoire variation (Preston, 1986, 1989, as cited by Clopper & Pisoni, 2005). The same listeners performed much better when broader labels were used to classify talkers. However, they were using geographic associations, and not ethnic ones, and the geographic regions involved tended to be in close proximity, unlike the typologically and culturally divergent ethnolinguistic groups used in this study.
Originally, it was planned to have a selection of talker photos that listeners could choose from to identify talker ethnicity as opposed to text labels. This was not realized, as many talkers did not consent to having their photos taken. Unfortunately, this phenomenon was observed more for the Chinese and Punjabi participants more so than for white participants. Therefore excluding those who did not want their photos taken might have resulted in a non-response bias. For example, perhaps the visible minority talkers were less comfortable with having their photos taken precisely due to the fact that they are considered visible minorities, even at the institutionalized level of Statistics Canada. As well, the study already suffers from significant non-response bias, as only those who responded to posters, announcements or mailing lists voluntarily. It was especially difficult to recruit Punjabi participants. In addition, while this strategy would have probably avoided the ambiguity and controversy surrounding ethnic/racial labels, and it could have also introduced potential confounds. For example, listeners might use features other than ethnicity to make their judgments, such as visual attractiveness.

Another limitation is that only sixteen listeners were used here. Although this means that ninety-six identifications for each talker were made, and nine hundred sixty for each ethnic group, the sample size may still not be sufficient to investigate other potential effects, such as gender, confidence, individual talker differences, or mother tongue, without compromising statistical integrity. This limits the inferences that we can make. However, more listeners have already been tested and we plan to continue recruitment until we have at least fifteen listeners with each network rank.

As well, no listeners had East Indian network ranks in our recruitment efforts at UBC. We plan to recruit more listeners, especially in areas where there are high concentrations of East Indians, namely Surrey, in order to look more at the interactions between familiarity and talker ethnicity identification. Judging from the patterns observed so far, it is probable that those with East Indian network ranks will be better at identifying East Indian talkers than listeners with other network ranks. I also suspect that identification accuracy may go up for East Indian talkers, and with more listeners in general, we will be able to discern more clearly whether listeners are more likely to misidentify East Indians as white or as Chinese.
Finally, a few other limitations must be mentioned quickly. First of all, only one speech style, which we dub here as ‘semi-spontaneous,’ was used here. This may not be the most representative, or control for all confounding variables. Therefore it might be useful to elicit data from different speaking styles, including conversations with different interlocutors. In addition, one white talker was an L1 French-English bilingual, who reported speaking French at home. Another Punjabi talker was raised in Terrace, B.C. According to Nayar (2012), Punjabis in this region exhibit a lesser degree of ethnic insularity than those in Metro Vancouver, and First Nations are an integral part of Terrace’s ethnic makeup. Further analyses will examine whether listeners exhibit unique identification patterns for these two talkers.

4.3 Future directions

Any differences between groups can unfortunately also serve as a source of discrimination. For example, Purnell, Idsardi, & Baugh (1999), using a matched guise technique, found results suggesting that housing discrimination can occur as a result of ethnically-associated differences in English. Thus one may argue that it is important to be aware of differences between groups, in that addressing such ‘linguistic profiling’ may be difficult if no one is explicitly aware them. Further research should be done to determine whether English usage differences between ethnic groups observed here can be used as sources of discrimination, which could be used in the law cases.

Another future direction is to look into the mechanisms behind the origins of these ethnolinguistic repertoires. As mentioned previously, the primary motivation for this project was the ongoing history of language contact between English, Punjabi and Cantonese. Nevertheless, just because contact was the motivation does not mean it is the mechanism. In addition, it is equally likely that more one mechanism may come into play here. As well, determining causation, statistically speaking, is near impossible in the social sciences, because human subjects are not bacteria and therefore random assignment is usually quite difficult to achieve.
In addition, useful theoretical models to apply to this situation might be borrowed from outside the field of linguistics, which include many possible sources of influence. For example, Bronfenbrenner’s (1979) bioecological model conceptualizes children’s social development as a set of nested contexts, each of which can exert bidirectional influence on each other (as cited by Siegler, DeLoache, & Eisenberg, 2011). An illustration of his model can be seen in Figure 5, below. The first level in which the child is embedded is the **microsystem**, which includes the child, and the activities, roles, and relationships with which they directly engage in over time. The next level is the **mesosystem**, which consists of the connections between different elements in the microsystem, such as a child’s immediate peers and their parents. The third level up is the **exosystem**, which consists of settings that the child may not directly participate but which still influence their development. These include mass media and parents’ workplaces. The next level, the **macrosystem**, consists of the broad ideology, laws, and customs of the individual’s culture, subculture, or social class. Finally, the highest level, the **chronosystem**, consists of the changes in the individual and their environment over time.
Thus the microsystem could be used to characterize the influence of parental language input, as well as an individual’s linguistic interactions with their peers. The mesosystem could illustrate how linguistic parental input and peer interactions might interact, and the exosystem could include the role of the mass media. The macrosystem might encompass elements such as the Canadian Multicultural Policy as well as the individuals’ heritage culture; the chronosystem might include how Canada’s immigration and voting laws have become more inclusive over time.

For example, due to the presence of Punjabi and Chinese ethnic enclaves in Metro Vancouver, Punjabi and Chinese Canadians have an opportunity to interact with peers
from the same ethnolinguistic backgrounds. Therefore their peers might be getting similar
linguistic input from both their parents and the surrounding exosystem and macrosystem,
that is input in both their heritage language as well as English. They may also receive
English input from their L2 English parents, who may not be fully proficient in English;
thus the children of first generation immigrants raised abroad may receive a very
different sort of English input, and may also have less access to the standard repertoire
(e.g. Clopper & Pisoni, 2004; Kerswill, 2010; Thomason, 2010). This might lead to code-
switching in Punjabi and Chinese Canadian peer-to-peer interactions in Metro
Vancouver, which in turn could lead to the establishment and stabilization of
ethnolinguistic repertoires. Finally, these individuals might transmit these repertoires to
their own children, who continue to use them and thereby modify the broader mainstream
culture of the macrosystem.

Despite the difficulty in establishing causal relationships, however, and leaving
aside theory for a moment, it may still be interesting to examine the relative contributions
of different potential mechanisms of change in Punjabi and Cantonese Canadian
ethnolinguistic repertoires. Language contact, transfer, speech communities, and identity
are several mechanisms that could be explored.

Correlations between linguistic features used in English by immigrant
descendants that are typical of their heritage language may imply that language transfer
was indeed the potential mechanism. For example, I have noted the usage of the Punjabi
word *panji* in English by Punjabi Canadians. *Panji* can be translated as ‘sister,’ and can
be used to refer to close female friends who are not related by blood as well.
Furthermore, Sharma and Sankaran (2011) have observed the usage of retroflex alveolar
stops in English among British-born Punjabis in Bradford, UK, a city of high Punjabi-
British concentration.

As well, various researchers have claimed to have found ‘rhythmic differences’
between English spoken by non-European ethnic groups whose languages have come into
close contact with English, including Cantonese (e.g. Lowe, Grabe, & Nolan, 2000; Mok
& Dellwo, 2008; Setter, 2006; Szakay, 2008; Torgersen & Szakay, 2011). This is
especially interesting when considering prenatal influence as a potential mechanism for
the establishment of an ethnolinguistic repertoire. That is, the womb is essentially a low-
pass filter, and beat frequency by definition is low frequency, meaning that much of prenatal input is largely suprasegmental.

Indeed, there is a wealth of evidence supporting the role of suprasegmental prenatal input on language acquisition. Cutler (2012) summarizes several of the perceptual studies with young infants. Using a high amplitude sucking procedure, Spence and DeCasper (1987) found results suggesting that newborns preferred unfamiliar voices in unfiltered form as opposed to low-pass filtered, but that they did not have a preference for either filtered or unfiltered forms of their mother’s speech. In addition, newborn infants appear able to discriminate between languages of different rhythm classes, such as Japanese and English, but not between languages within the same rhythm class, such as Dutch and English (Nazzi, Bertoncini, & Mehler, 1998). Moreover, Setter (2006) summarizes evidence suggesting that non-native speakers may have a hard time understanding a second language when it is of a different rhythm class than their own, and that native-like speech rhythm may one of the hardest things for second language learners may acquire. This leads one to speculate that Cantonese and Punjabi Canadians may adopt the suprasegmentals of Cantonese or Punjabi while speaking English to members of their ethnolinguistic community who have poor English proficiency. Additionally, the suprasegmentals of near-native English speakers’ Cantonese and Punjabi L1s may have been so ingrained in them that despite being English-dominant they may continue to use them while speaking English. Therefore, heritage language maternal input may have a way to seep into the linguistic resources of even in English-dominant Chinese and Punjabi Canadians who are unable to hold a conversation in their heritage language. However, nobody seems to agree on how to measure speech rhythm let alone its definition, and few researchers have used perceptual studies in order to classify rhythmic typology between or within languages, despite mounting evidence that isochrony is not common in production (Kohler, 2008).

Also, I recently noticed paper signs were observed taped to stalls in the women’s bathrooms at the Koerner Library at the University of British Columbia reading “Please do not throw napkin or paper towel in the toilet.” Generally, most standard varieties of English mark pluralization morphologically on nouns. On the other hand, Chinese
languages lack plural morphology (Lai-Shen Cheng & Sybesma, 1999), and lack of plural morphology has also been seen in Hong Kong English (Gisborne, 2011).

While it is not clear whether a Chinese person created this sign, lack of morphological noun pluralization using –s on English language signage can be seen in Chinese-owned businesses in other areas of Metro Vancouver as well, particularly in places such as Richmond or Metrotown in Burnaby where the concentration of Chinese people is relatively high. Another counterpoint is that perhaps it is only L2 English Chinese who are writing such signs. However, perhaps the presence of such signs still serves to reinforce English with Chinese-style pluralization in Metro Vancouver among L1 English Chinese who go to such shops. In any case, it is necessary to evidence my anecdotal observations with appropriate documentation and frequency estimates.

Likewise, this goes for other informal observations I have noted here, such as the use of tūshēng huáyì /tou2 saang1 waa6 jeoi6. Is the usage reported by the Canadian-raised Cantonese talker and China-raised Mandarin talker typical of the overall Metro Vancouver communities of Canadian-raised Cantonese-speaking Chinese and China-raised Mandarin talkers, or are the anecdotal observations here just idiosyncratic of the two particular talkers?

Moreover, in order to confirm the importance of the role of familiarity in identifying these talkers’ ethnicities, further research might include listeners from outside of Metro Vancouver. As well, it would be insightful to see whether these ethnolinguistic repertoires are specific to Vancouver in terms of production.

In addition, although results do not support the existence of a general Asian English repertoire in Vancouver, it is possible that there may be generalized East Asian and South Asian repertoires. For example, Koreans also make up a fair portion of Metro Vancouver’s population, and there are cultural and typological similarities between Korean and Chinese. For instance, Korean has many loanwords from Chinese languages (Matthews, 2010). In addition, Koreans and Chinese both make up a large proportion of the city of Burnaby within Metro Vancouver (Statistics Canada, 2012a). However, ethnic tension exists between the two groups, and Korean grammar differs markedly from Chinese. The former notion can be reflected in the joseonjok, or Koreans in China and the racism they experience in both South Korea and China (Lee, 2009).
Some *joseonjok* simply do not know whether they should identify themselves as Chinese or Korean. Thus, in many ways, *joseonjok* experience a similar identity crisis to that of Chinese and Punjabi Canadians, and it would be interesting to find out if Korean and Chinese Canadians shared a similar English repertoire. This is in part because a shared repertoire might be a positive sign for Korean-Chinese relations and discourage negative views associated with the *joseonjok* in South Korea.

It would also be very interesting to examine whether there is even a generalized Chinese ethnolinguistic repertoire in Vancouver English; that is, research exploring the possibility that Chinese Canadians of multiple heritage language backgrounds may speak English similarly. This inquiry is especially intriguing, like Korean and Chinese, there are multiple similarities between these groups, but also several differences, and enough cultural tension to make divergence also likely.

Another research direction involves the question as to whether the visible minority groups participate in regional changes in progress, such as the Canadian Shift and Canadian raising. Hoffman and Walker (2010) report no significant differences between Chinese, Italian, and Anglo Canadians in Toronto with respect to F1 and F2 for Canadian Shift vowels. Currently I am helping Dollinger (forthcoming) analyze the Canadian Shift and Canadian raising for Vancouverites and Washingtonians among different age groups and between Chinese and Anglo Canadians. Results may either support or challenge the notion of Canadian English “autonomy” from American English, as it has been claimed that the Canadian Shift is a shared phenomenon across Canadian regions and the most consistent characteristic of contemporary Canadian English (Dollinger & Clarke, 2012).

Finally, there is no reason to assume that the language transfer that could be involved here is bidirectional. A possible study could involve recording Cantonese Canadians and Punjabi Canadians talkers in addition to talkers raised in Hong Kong, Guangdong, and the Punjab provinces of Pakistan and India, while they speak Cantonese and Punjabi. We could then present them to listeners and find out if perhaps Cantonese and Punjabi Canadians also employ unique ‘Canadian English’ ethnolinguistic repertoires. It might even be possible to record these talkers in English as well, and find
out if there is an ethnolinguistic repertoire that is commonly shared between languages, to
examine theoretical questions about bilingualism and interlanguages.
5 Conclusion

Listeners’ ability to identify the ethnicities of talkers from three different ethnic backgrounds suggests the presence of differences in the way each group speaks English. Although the cause and nature of these differences are unclear, the results of this study provide a good starting point for finding the features of these ethnolinguistic repertoires and linking them to broader theories of language contact, ethnolinguistic identity, and speech communities.

In addition, the relationship between listener network rank and identification accuracy emphasizes the roles of familiarity and experience. Moreover, taken together with listeners’ misidentification patterns, it appears that ethnically marked variation in English may be more accessible to in-group members. Thus Chinese and East Indian Canadians may have found their own ‘safe space’ in Vancouver after decades of struggling for identity in the face of decades of discrimination.
References


Appendix A: *The North Wind and the Sun*

The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.
Appendix B: Talker background questionnaire

Language Background Questionnaire

- Are you male or female? ______________
- What is your race (check all that apply):
  - First Nations
  - Asian
  - Pacific Islander
  - Black
  - White
  - Hispanic
  - South Asian
- What is your age? ______
- What cities or towns have you lived in? List first the place where you were born, and list each town or city you have lived in.
  
  birth until age ___ in town/city _____________________
  age ___ until age ___ in town/city _____________________
  age ___ until age ___ in town/city _____________________
  age ___ until age ___ in town/city _____________________
  age ___ until age ___ in town/city _____________________

- What languages do you speak (include your native language(s))? When did you start learning this language? How would you rate your proficiency in reading, writing, speaking, and understanding it? (1) not at all, (2) poorly, (3) fairly well, (4) fluently.

<table>
<thead>
<tr>
<th>language</th>
<th>age</th>
<th>reading</th>
<th>writing</th>
<th>speaking</th>
<th>understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>__________</td>
<td>___</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>__________</td>
<td>___</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>__________</td>
<td>___</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>__________</td>
<td>___</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

- Which language(s) do you most commonly speak….
  At home? __________ At work? ____________ At school? ____________
  With friends? ____________

- Do you have any speech or hearing disorders?
  If “yes”, please specify. no yes

- Are you right-handed or left-handed?

- What is the highest level of education that you have completed?

- What is the highest level of education that your parents have completed?
• What are the professions of your parents or other caregivers?

• Where were your parents born and raised?

• Where were your grandparents born and raised?
Appendix C: Listener Background Questionnaire

Language Background Questionnaire

Are you male or female? ______________

• What is your race (check all that apply):
  ___First Nations  ___Asian  ___Pacific Islander  ___Black
  ___White  ___Hispanic  ___South Asian

• What is your age? _____

• What cities or towns have you lived in? List first the place where you were born, and list each town or city you have lived in.

  birth until age __ in town/city _____________________
  age __ until age __ in town/city _____________________
  age __ until age __ in town/city _____________________
  age __ until age __ in town/city _____________________
  age __ until age __ in town/city _____________________

• What languages do you speak (include your native language(s))? When did you start learning this language? How would you rate your proficiency in reading, writing, speaking, and understanding it? (1) not at all, (2) poorly, (3) fairly well, (4) fluently.

<table>
<thead>
<tr>
<th>language</th>
<th>age</th>
<th>reading</th>
<th>writing</th>
<th>speaking</th>
<th>understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>__________</td>
<td>__</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>__________</td>
<td>__</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>__________</td>
<td>__</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>__________</td>
<td>__</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

• Which language(s) do you most commonly speak….
  At home? ______________ At work? ______________ At school? ______________
  With friends? ______________

• Do you have any speech or hearing disorders?
  If “yes”, please specify.  no yes

• Are you right-handed or left-handed?

• What is the highest level of education that you have completed?
  ________________________

• What is the highest level of education that your parents have completed?
  ________________________
• What are the professions of your parents or other caregivers?

_________________________________________________________________

• Where were your parents born and raised?

_________________________________________________________________

• Where were your grandparents born and raised?

_________________________________________________________________

• Think about your friends, family and coworkers. With which of the following ethnic groups do you spend the most time? Please rank them in order, from 1 for the group with which you spend the most time, to 3 for the group with which you spend the least time.
  __ East Indian
  __ White
  __ Chinese