NIAGARA ENGLISH:
LANGUAGE VARIATION AND DIFFUSION ON THE U.S.-CANADA BORDER

Claire Henderson

Department of Linguistics
McGill University, Montreal

August 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts.

Table of Contents

Abstract (English)

This thesis examines patterns of linguistic diffusion in the context of a national border, looking at the potential spread of Canadian and American English variants between the Niagara border regions of Ontario and New York. Using (1) a dialect questionnaire and (2) acoustic analysis of sociolinguistic interviews, I outline patterns of diffusion and non-diffusion. The results show that the border primarily acts as a barrier but that certain variables do show potential wavelike and hierarchical diffusion. In particular, the findings suggest that phonemic incidence variables (e.g., foreign (a) pronunciation) are more likely to diffuse than structural variables (e.g., vowel shifts) and Canadian markers (e.g., vocabulary, spelling). Additionally, the border shows a weakening effect for instances of diffusion. In light of these findings, I discuss the role of the border in models of diffusion as well as identity, stability, and change in Canadian English.

Abstract (French)

Le but de cette thèse est d'étudier les modes de diffusion linguistique en contexte d'une frontière nationale, ciblant la diffusion possible de variantes linguistiques en anglais canadien et américain entre la région de la frontière Niagara en Ontario et celle en New York. En employant (1) un questionnaire dialectal et (2) l'analyse acoustique d'entretiens sociolinguistiques, je décris les cas de diffusion et non-diffusion. Les résultats démontrent que la frontière agit d'abord comme une barrière pour la diffusion, mais qu'effectivement certaines variables montrent la diffusion par vague et la diffusion hiérarchique. Notamment, les résultats proposent que les variables d'occurrence phonémique (ex. la prononciation de « foreign (a) ») ont plus de tendance à diffuser que les variables structurales (ex. décalages de voyelles) et les marqueurs d'identité canadienne (ex. vocabulaire, orthographie). La frontière provoque aussi un effet d'affaiblissement pour les cas de diffusion. Étant donné tous ces résultats, je présente une discussion du rôle de la frontière pour les modèles de diffusion ainsi que l'identité, la stabilité et le changement en anglais canadien.

Acknowledgments

I am extremely grateful to everyone who participated in these studies for their time and their enthusiasm for this project. This would not have been possible without them.

I am also very thankful for all of the insight and advice I have received from my supervisor Charles Boberg in developing and implementing both studies. He has always been encouraging and extremely helpful throughout my work on this and I know it would not be the same without his help. Thank you also to my committee member Morgan Sonderegger for his helpful advice on acoustic measures and data analysis.

I would also like to thank my classmates and the audiences of the 2023 American Dialect Society Annual Meeting and the 2023 Canadian Linguistics Association Annual Meeting for their comments which have helped shaped my approach. I am especially grateful to Jeanne Brown for her support and advice and to Irene Smith and Massimo Lipari for their guidance and patience teaching me about acoustic measures in R.

Thank you as well to all of my family and friends for their support. Special thanks to Nathan for his encouragement and for pronouncing all of the unmerged LOT words for me. I am also grateful to my cat Lady for her company and commentary.

This research is supported in part by funding from the Social Sciences and Humanities Research Council and funding from the Fonds de recherche du Québec.

List of Figures and Tables

Figures:

Tables:

List of Abbreviations

F1 – first formant

F2 – second formant

GTA – Greater Toronto Area

LBMS – Low-Back-Merger-Shift

NARVS – North American Regional Vocabulary Survey

NCS – Northern Cities Shift

NNY – Niagara, New York

NON – Niagara, Ontario

NY – New York (State)

ON – Ontario

SD – standard deviation

U.S. – United States

1. Introduction

1.1 General introduction

The diffusion of linguistic innovations across speech communities is a fundamental part of variation and change. As features spread from one community to another, those speakers come to sound more similar to each other. However, where a given feature diffuses and whether it becomes adopted by speakers depends on many factors. Previous research shows that diffusion largely depends on the size and distance of speech communities with two main patterns: wavelike and hierarchical. Wavelike diffusion involves linguistic features spreading in waves from one centre of innovation to nearby areas, thereby depending on the distance between communities (Labov 2007). Hierarchical diffusion instead involves spread between larger centres before the features reach smaller communities in between and thus depends more on community size (Trudgill 1974, Callary 1975). The actual pattern varies based on the particular features and communities, as well as social factors and attitudes (Bailey et al. 1993).

National borders like that between Canada and the United States present a particularly interesting opportunity for studying diffusion. We might expect border towns to become very linguistically similar to each other due to proximity and contact between speakers. However, the border represents a socio-political boundary which may limit diffusion, especially if the feature is perceived as a marker of national or local identity. Previous work has shown the border to primarily act as a linguistic barrier but diffusion of certain variables has also been found (e.g., Allen 1959, Miller 1989, Boberg 2000, Burnett 2016). In this thesis, I investigate potential border effects on language use near the Niagara border between Southern Ontario and Upstate New York. Specifically, I look at patterns of diffusion and non-diffusion for a wide range of variables, focusing primarily on pronunciation.

The Niagara Region in Canada largely consists of small towns and cities, with three major border crossings to the United States. Niagara is at the bottom tip of the Golden Horseshoe area of Southern Ontario, which stretches along Lake Ontario up to the Greater Toronto Area. On the American side, Niagara includes Buffalo as well as smaller towns along the Niagara River. Niagara, NY is part of the Inland North speech area of the U.S., characterized by the Northern Cities Shift (Labov et al. 2006).

To investigate the effect of the border, I will look at three regions: Niagara, ON; Niagara, NY; and the Greater Toronto Area (GTA). By comparing the three regions, it is possible to differentiate potential patterns of diffusion. Wavelike diffusion would predict greater use of American forms in Niagara, ON such that the three regions form a dialect continuum. Hierarchical diffusion would predict greater use of American forms in the GTA since Toronto is a much bigger speech community than Niagara. On the other hand, it is possible that both Canadian regions will be much more similar to each other and equally different from New York, indicating a strong border effect.

The primary research question is therefore whether the border has an effect on diffusion: do Canadian-American differences remain distinct near the Niagara border or does any type of diffusion occur? Additionally, my goal was to determine which types of features are more or less likely to diffuse. To answer these questions, I use both a questionnaire study and an acoustic study. The questionnaire study focuses on the types of features speakers can more easily report themselves: vocabulary, low-level grammar, spelling, and certain pronunciation variables. The acoustic study expands on pronunciation by looking at phonetic and phonological differences such as vowel shifts and Canadian Raising. In both studies, I compare the use of variants in the three regions to establish diffusion patterns and more generally examine the differences, similarities, and potential changes in American and Canadian English.

The rest of this chapter will provide the necessary background on diffusion and Canadian English (1.2). Chapter 2 will outline the methodology, results, and analysis of the questionnaire while Chapter 3 will outline the acoustic study. Chapter 4 presents a combined discussion of the two studies, focusing on the overall findings for diffusion, non-diffusion, and Canadian-American differences. Chapter 5 summarizes these findings and concludes on the main research questions.

1.2 Literature review

1.2.1 Diffusion

The question of how dialects diverge and converge has long been a foundational issue of dialectology and sound change. Bloomfield (1933) noted in particular that the spread of linguistic innovations depends largely on who speakers directly communicate with and is inhibited by breaks in the network of communication, including both geographical and social boundaries. Diffusion describes this type of change, which spreads between different dialect communities through contact as opposed to community-internal transmission over generations (Labov 2007). Diffusion

is thus largely a result of contact between adult speakers, meaning that structural patterns which are more difficult for adults to learn are less likely to diffuse (Labov 2007). While transmission leads to divergence as each community changes independently, diffusion leads to convergence as more communities share the diffused features.

There are two basic views on the process of diffusion: the wave model and the hierarchical model. The wave model assumes linguistic features spread from centres of innovation through contact with other nearby speech communities, leading to a wavelike pattern of diffusion from these centres, in line with Labov's (2007) diffusion data and observations. The degree to which a community uses an innovative form is then a function of how far the community is from the centre of innovation. In contrast, the hierarchical model assumes the main factor in diffusion is community population size: innovations spread from large urban communities to other large communities before reaching the smaller communities in between (Trudgill 1974, Callary 1975).

Trudgill (1974) introduced further methodology in the study of linguistic geography and diffusion, drawing from a combination of contemporary sociolinguistic theory and theoretical geography. He argued that being able to actually quantify geolinguistic diffusion would allow us to both describe and explain the distribution of features. Noting that language diffusion is similar to other types of diffusion, such as the spread of technological innovations, he proposed a linguistic gravity model based on gravity models used by geographers. The model considers distance between dialect areas and their population size, such that linguistic innovations are expected to spread between larger nearby cities before infiltrating the smaller, less populated areas. It is therefore largely rooted in the hierarchical view of diffusion. The model also includes a linguistic similarity index to account for languages or dialects which are generally more similar than others.

However, this model has been criticized for excluding potentially relevant factors for diffusion. The most apparent gap is that its assumption of hierarchical diffusion does not account for the wavelike diffusion that can be found for certain variables. This is noted by Bailey et al. (1993) who looked at apparent time diffusion of grammatical, phonological, and lexical features in Oklahoma. They found patterns of hierarchical and wavelike ("contagious") diffusion as well as instances of combined hierarchical and wave diffusion (e.g., laxing of /u/ in *pool* spreading west to east while also following an urban-rural hierarchy) and instances of contrahierarchical diffusion (e.g., use of *fixin' to* spreading first through rural communities). The gravity model does not predict

these other patterns of diffusion and therefore cannot capture the spread of non-hierarchical variables.

Another gap in the gravity model is that by simplifying the key causes of diffusion, it misses other important variables that may inhibit or enhance diffusion. This is also emphasized by Bailey et al. (1993) who highlight the importance of social and attitudinal influences, particularly in relation to the external or internal prestige associated with the feature. For example, they show that external prestige variables like unrounding of the vowel in *hawk* spread hierarchically while internal prestige variables like use of *fixin' to* spread contrahierarchically. A model based largely on geographical factors cannot account for these interactions with speaker attitudes.

Similarly, Boberg (2000) points out that Trudgill's model does not consider the effects of national boundaries. This is a particularly important consideration for neighbouring countries with high linguistic similarity, such as the United States and Canada. By looking at different types of pronunciation variables in Windsor and Detroit, Boberg's (2000) study showed that the border functions as a barrier to diffusion for some variables but not others. Notably, the extent and method of diffusion varied by the type of linguistic feature being examined. While the border acted as a strong barrier for the phonemic inventory measure, the pronunciation of foreign (a) variables showed wave diffusion across the border with the Windsor speakers near the border having more American pronunciations than Toronto speakers. This suggests that some wave diffusion is possible for non-structural features (e.g., phonemic incidence) while structural features (e.g., phonemic inventory) tend to be more clearly divided between Canadian and American English. In either case, the hierarchical gravity model does not account for the attested patterns. The role of the national border must therefore be considered in the description and explanation of geolinguistic diffusion.

Diffusion near the U.S.-Canada border has been of particular interest for many years. In the 1950s, Avis (1954, 1955, 1956) investigated differences on the Ontario-U.S. border. He suggests that the settlement history of the Ontario border contributes to language use on both sides: the early Loyalist settlers from the U.S. and the interaction between both sides of the border likely meant that American speech had a greater influence in these areas. Differences then may have arisen from a combination of British immigration, innovations, and varying outcomes for competing forms. For pronunciation in particular, he found that preferred variants tend to be

British “when the word is literary and consequently of limited currency in speech” (Avis 1956: 55) and American “when the word is in widespread general use” (Avis 1956: 55).

The question of whether the border acts as a barrier was then addressed in Allen’s (1959) study of the middle border region between the U.S. (Minnesota, North Dakota) and Canada (Saskatchewan, Manitoba, Ontario). He notes many distinctive Canadian features that do not diffuse, but also identifies nearly 50 Canadian English variants found in the American border regions. These include several vocabulary items and some phonological features, including Canadian Raising. At the same time, there are few American variants that did not also appear in the Canadian regions. He therefore suggests that while diffusion does occur in both directions, American features are more readily adopted by Canadians than vice versa, meaning that the border acts more as a barrier for the Americans.

This topic was further investigated along the Maine-New Brunswick border by Miller (1989) and Burnett (2016). Miller (1989) found both overlap and separation between the speakers on either side of the border for vocabulary and pronunciation forms, concluding overall that the border constitutes a linguistic boundary. One important variable he notes is the split in /ɑ/ vs. /æ/ pronunciations in words like *drama*, *pasta*, and *Mazda* (i.e., foreign (a)), with Canadians overwhelmingly preferring /æ/. He specifically emphasizes the role of language in identity and the Canadians’ need to separate themselves from the Americans. Burnett (2016) found that rather than converging towards American forms, the Canadians were primarily maintaining and even increasing use for many of the Canadian variants studied. The cross-border influence that did occur was found in both directions. She also links these findings to identity and autonomy: the border separates speakers who are otherwise in a dialect continuum and the speakers tend to act more like their own group (“Canadian” or “American”) in terms of language and other social behaviours. More generally, these two studies suggest that the border mainly acts as a barrier and that Canadian English is being actively maintained by speakers as an identity marker.

Swan (2020) looked at the linguistic effect of the border in relation to individual views and background. She found that the raising and fronting of /æ/ before /g/ (BAG-raising) in the Seattle-Vancouver area was linked to ideological stances on local issues: those who were more concerned about encroachment and change within their city were more likely to raise. This parallel finding in both cities shows a connection between BAG-raising and local identity, highlighting the

importance of individual identity and ties in the use of features perceived as local markers, particularly for speakers close to an international border.

There is limited work on the Niagara border specifically, but the region was included in Chambers' (1994) Dialect Topography project. This project collected questionnaire responses throughout the Golden Horseshoe region of Southern Ontario and the nearby areas in New York State. Chambers (1994) identifies some clear boundaries at the border, such as the pronunciation of *shone* which turns from 95% [ʃon] in the U.S. to 92% [ʃɑn] in Canada. He also notes a general tendency for Canadians to accept more lexical variation (e.g., several variants for *wedgie*) than Americans (e.g., uniform use of *wedgie*).

Using the Dialect Topography data, Easson (1999) specifically examined Canadian-American differences around the Niagara border. He found two variables (*couch* vs. *chesterfield* and yod-dropping in *news/student*) for which Canadians are becoming more similar to Americans, but many more variables that remain distinct throughout the Golden Horseshoe region. For one variable (pronunciation of *lever*), he shows that Niagara Falls has the same frequency of the Canadian variant as the rest of the region and that there is a sharp drop-off across the border, suggesting little cross-border influence in Niagara. On the other hand, he finds that Canadian *running shoes* (rather than American *sneakers*) has lower use in Niagara than in the North Shore of the Golden Horseshoe suggesting some wavelike adoption of the American form. However, there is still a large drop-off across the border and the use of the Canadian variant increased in apparent time. Overall, this study provides further evidence that the Niagara border acts mainly as a barrier to diffusion and that speakers maintain their Canadian English use and associated identity.

1.2.2 Canadian English

In order to look at cross-border diffusion we must first establish what features characterize Canadian English, particularly in relation to American English. As noted above, many differences have been identified in vocabulary, grammar, spelling, and pronunciation. There have been a number of studies on Canadian English features using both questionnaires and acoustic analysis. Here, I will focus on the major questionnaire studies that form the foundation of my own questionnaire and outline previous work on the key phonetic variables I will examine in the acoustic study.

There have been many questionnaire-style studies of Canadian English over the years. These studies target features that are salient enough for speakers to identify themselves, such as

vocabulary and phonemic incidence. As described in 1.2.1, Avis (1954, 1955, 1956) conducted several investigations via questionnaire of the Ontario-U.S. border to establish differences which separated Canadian English from the rest of the northern speech area in the U.S. These studies report several distinctive vocabulary (e.g., *blinds*, *tap*, *serviette*), grammar (e.g., *lend*, *quarter to*), and pronunciation (e.g., *vase* /vez/, *been* /bin/, *produce* with /ɑ/) variables. Avis primarily links differences to British influence and similarities to the Loyalist settlement history with little independent Canadian innovation.

The study of Canadian English then expanded greatly with the Survey of Canadian English (Scargill & Warkentyne 1972). This survey included every province and gathered responses across two generations (ninth graders and their parents). The authors establish trends across the two generations and between the different provinces, creating a more complete picture of Canadian English. The questionnaire included over 100 questions on vocabulary, pronunciation, grammatical usage, and spelling. They show for example that Canadians say *zed* over *zee*, rhyme *lever* with *beaver*, and rhyme *cot* with *caught*. They also show that Canadians have several vocabulary differences such as *tap* over *faucet*, *chesterfield* over *sofa*, and *icing* over *frosting*. Canadians also had a strong preference for certain British spellings like *travelled* and *defence*. The grammatical usage variables pattern less clearly and often show split usage, though there is a big change across the generations from *sneaked* to *snuck* (as the past tense of *sneak*). In general, this study makes it clear that Canadian English stands apart from American English in many ways and that the degree of American or British influence largely depends on the variable.

The next major questionnaire study was J.K. Chambers' Dialect Topography project in the 1990s, which primarily focused on the Golden Horseshoe region in Ontario. The data showed further vocabulary and pronunciation differences that set Canadian English apart, such as the pronunciation of *asphalt* with /ʃ/, and reaffirmed several variables such as *zed* (vs. *zee*) (Chambers 1994). There was also some decline in Canadian variants such as the change from *chesterfield* to *couch* (Chambers 1995). Since the data are separated by smaller regions rather than provinces, some of these studies have also shown differences within Canada and patterns of diffusion, as in Easson's (1999) analysis of the Niagara border data (see 1.2.1).

Over the years, many of the vocabulary variables have become outdated, such as technical farming terms. Boberg (2005) introduced many modern vocabulary variables in the North American Regional Vocabulary Survey (NARVS). This questionnaire collected response across

Canada and the United States to identify key vocabulary differences between and within the two countries. Several strong Canadian variants were found, including *grade one* (vs. *first grade*), *chocolate bar* (vs. *candy bar*), *bachelor apartment* (vs. *studio apartment*), *bank machine* (vs. *ATM*), and *washroom* (vs. *restroom*). Many older variables also showed a persistent strong division, such as *tap* (vs. *faucet*), *zed* (vs. *zee*), and *eavestroughs* (vs. *gutters*). NARVS therefore demonstrates the continuous role of lexical variation in how Canadian English separates itself from American English.

As mentioned in 1.2.1, Boberg (2000, 2009, 2020) also looked at the pronunciation of foreign (a) (orthographic <a> in various loanwords) in Canada. These variables (*pasta*, *drama*, *taco*, *pajamas*, etc.) have traditionally favoured the /æ/ pronunciation in Canada and /ɑ/ in the U.S. However, Boberg (2000) showed that the American pronunciation is gradually diffusing to Canada and further acoustic analysis confirms that it continues to diffuse (Boberg 2009, 2020). A potential intermediate pronunciation is also identified in some cases, though its use has proven relatively rare (Boberg 2009, 2020). This change suggests that not all variables continue to show a strong division; we must therefore consider the type of variable being studied when investigating cross-border differences and diffusion.

The type of variable is particularly important for variation in "pronunciation." While the differences mentioned thus far have largely dealt with phonemic incidence (i.e., different lexical distributions of the phonemes within the speaker's system), there are also a number of phonetic and phonological differences. In the acoustic study, I will look at several of these differences in addition to foreign (a): the low back merger and the Northern Cities Shift, the Low-Back-Merger Shift, Canadian Raising, allophones of /æ/, and /u:/-fronting (see Appendix B for vowel notation and example words).

First, the Northern Cities Shift (NCS) is important to the study of Niagara since the American side of the border participates in the shift. The primary components of the NCS are the fronting of the LOT vowel (/ɑ/) and the tensing of the TRAP vowel (/æ/), leading to raising and fronting (Labov 1991). These movements are in part due to the non-merger of the LOT and THOUGHT (/ɔ/) vowels which leads to more crowding among the low vowels; LOT therefore moves forward while TRAP moves upwards along the front periphery (Labov 1991). In contrast, Canadians overwhelmingly have the low back merger and so LOT-THOUGHT remains back and TRAP remains

low. We therefore expect the Canadians to show the merger and no NCS regardless of distance from the border as these structural features do not typically show diffusion (Labov 2007).

On the other hand, the Low-Back-Merger Shift (LBMS) should only affect the Canadians since it depends on speakers having the low back merger. In the Atlas of North American English, Labov et al. (2006) consider this shift to be a defining feature of Canadian English, distinguishing it from nearby varieties of American English. The main component of the shift is the retraction and/or lowering of three short front vowels: KIT (/ɪ/), DRESS (/ɛ/), and TRAP (/æ/) (Clarke et al. 1995, Labov et al. 2006). Back vowel movement is also sometimes linked to the LBMS, though the pattern varies across findings and regions (Boberg 2021). There is also apparent time evidence suggesting this shift is a change in progress with younger speakers having more retraction and/or lowering (Labov et al. 2006; Boberg 2008, 2010, 2019a, 2019b). Women also tend to be ahead in this shift (Clarke et al. 1995, Boberg 2010). It should be noted that the NCS also involves backing and/or lowering of KIT and DRESS (Labov 1991), which means the two shifts will primarily be differentiated by the behaviour of TRAP and LOT.

Another major feature of Canadian English is Canadian Raising. This phonological pattern involves raising of the nucleus of the diphthongs /ai/ and /au/ before voiceless obstruents; for example, /ai/ in *knife* is expected to raise while /ai/ in *knives* stays in place (Chambers 1973, Boberg 2010). Canadian Raising entails not only a change in height (F1) but also a change in backness (F2). For /ai/, the nucleus is raised and fronted (Boberg 2010). The change in backness for /au/ is more regional; in Ontario, /au/ is raised and fronted (Boberg 2010). This means that the Canadians in Niagara and Toronto are expected to show raising and fronting of both /ai/ and /au/. However, this does not mean the Americans do not have Canadian Raising: /ai/-raising has been frequently reported in the Inland North (i.e., including Upstate New York) (Boberg 2010). On the other hand, /au/-raising is less widespread though it has been reported in Martha's Vineyard, Boston, Virginia, and New Orleans (Labov 1972, Chambers 1973, Boberg 2010, Carmichael 2020). The Niagara, NY group in the current study would therefore be expected to show /ai/-raising only, unless there is diffusion from Canada or elsewhere.

The next acoustic feature that will be examined deals with allophones of /æ/ before voiced velars (/æG/; G = /g/, /ŋ/) and front nasals (/æN/; N = /n/, /m/). In Canada, both of these allophones tend to be raised and fronted in comparison to /æ/ in other environments, with /æN/ being the more advanced allophone in Ontario (Boberg 2008, 2010). As with the LBMS and Canadian Raising,

this was considered an identifiable feature in Canadian English, though it is again not limited to Canada (Labov et al. 2006). The relative distance from /æN/ and /æG/ to /æ/ also relates to the merger and shifts described above: the low, retracted /æ/ in Canada means it has a large difference with the raised allophones. In contrast, the tensing of /æ/ in the NCS means the Americans raise and front all of the /æ/ allophones and so all three are relatively close together, though Labov et al. (2006) also note that the pre-nasal allophone tends to be the highest and frontest of this group.

Finally, the fronting of /u:/ (GOOSE) is an important variable more generally in North America, varying across regions. In particular, /u:/-fronting was found to be more advanced in Canada but more conservative in the Inland North (Labov et al. 2006; Boberg 2010, 2021). This fronting applies to /u:/ when it is not followed by /l/. It is primarily a result of phonetic pressures on back vowels and the maintenance of contrast, occurring frequently cross-linguistically (Hock 2021). Along with the retraction of /æ/, this feature gives a general view of how innovative or conversative speakers are in the North American context (Boberg 2021). The relative movements of /æ/ and /u:/ are therefore combined in the Index of Phonetic Innovation (IPI) as a measure of innovation (Boberg 2010, 2021). Canadians have been shown to have more negative, innovative scores, particularly in Southern Ontario (Boberg 2010:204), whereas Americans in the Inland North tend to have positive, less innovative scores (Boberg 2021:225). This reflects the higher degree of both /æ/-retraction and /u:/-fronting in Canada.

Taken together, these various acoustic measures help to separate the expected Canadian and American patterns in the regions of study. More generally, many of these previous findings show strong Canadian-American differences for both the acoustic and questionnaire variables. However, some clearly have shown diffusion such as foreign (a) and other phonemic incidence and vocabulary items. In comparison to the questionnaire results, it seems that acoustic vowel differences do not tend to diffuse. Looking at all of these patterns near the Niagara border will help further determine whether diffusion occurs and which types of variables are more likely to diffuse.

2. Questionnaire study

2.1 Introduction

The first investigation into diffusion near the Niagara border involved a dialect questionnaire. Participants were recruited from the three key regions: Niagara, ON; Niagara, NY; and the GTA. Based on the previous research on the border (reviewed in 1.2), I expected the border to primarily be a barrier to diffusion. Given that the previous findings have linked the barrier effect to the maintenance of Canadian identity, we may expect that the types of variables which show less diffusion would be more identifiably "Canadian." For example, vocabulary may be particularly recognizable and thus less likely to diffuse if speakers are trying to maintain local identity. Similarly, the spelling variables are generally learned in school making them more likely to be recognizable as "Canadian" as well as more ingrained in the speaker's linguistic system. I will show that this is in fact the case for both vocabulary and spelling as they show little diffusion.

However, since the questionnaire deals with non-structural variables it is possible that some features will diffuse; for example, I expected to see wavelike diffusion of foreign (a) variants as was found for the Detroit-Windsor border (Boberg 2000). I also generally expected the pronunciation variables to diffuse more easily across the border since these features are likely to arise in interactions between speakers. They may also be less identifiable as "Canadian" and thus less linked to identity. Additionally, most of the grammatical variables examined here have shown similar usage on both sides of the border in previous studies (e.g., Avis 1956, Scargill & Warkentyne 1972) and are therefore expected to show no border effect at this point. However, there is one newer variable that was not included in the previous questionnaires which I expected to show a cross-border difference (*done (with)*). The results will show that pronunciation is the most likely to diffuse, but that the pattern is dependent on the variable. Similarly, grammatical diffusion depends largely on the variable with only *done (with)* forming a strong barrier.

This chapter will describe the questionnaire study and its findings. In section 2.2, I outline the methodology used. In 2.3, I detail the results of the study, looking at overall trends and specific variables of interest. In 2.4, I analyze these findings in relation to the research questions.

2.2 Methodology

The main goal of the questionnaire was to differentiate between Canadian English and American English variant usage for notable vocabulary, pronunciation, grammar, and spelling

variables. These variables were largely drawn from previous studies: Avis's (1954, 1955, 1956) studies of the Ontario-U.S. border, Scargill and Warkentyne's (1972) Survey of Canadian English, Chambers' (1994, 1995) Dialect Topography project, Boberg's (2005) survey of modern Canadian English vocabulary, and Boberg's (2000, 2009, 2020) studies of foreign (a) pronunciation. In addition to language data and demographic information, the questionnaire also collected ideological views related to Canada and the United States. The data were then analyzed according to region as well as linguistic and social factors. In this section, I will outline the methodology used for the questionnaire study including the participants and data collection (2.2.1), the questionnaire used (2.2.2), and the data analysis methods (2.2.3).

2.2.1 Data collection

Participants for the questionnaire were native English speakers aged 18 years or older who grew up in one of the three regions of study. The initial goal was to collect at least 40 responses from each region. Participants who were currently living in a different city were accepted unless they had moved to one of the other two regions being studied. Though people who have moved are perhaps not the ideal participants for regional dialect studies, they do reflect the true nature of the regions' populations, particularly in border towns. Additionally, the variables were mostly dependent on Canadian and American differences rather than regional differences within each country that could be more heavily affected by migration.

To gather a large quantity of data in a short time period, an online questionnaire was developed and distributed through social media platforms. The questionnaire was open for approximately three months in early 2022. No incentive was offered. In total, 803 usable responses were collected. The distribution of these responses will be outlined in the results (2.3.1).

2.2.2 The questionnaire

The questionnaire used in this study is largely based on previous studies investigating Canadian and American dialect differences. Many questions were reproduced in their original form to allow for a comparison with expected Canadian and American features, though a few new questions and answer options were added. A large proportion of the variables were derived from Scargill & Warkentyne's (1972) survey of Canadian English. These include pronunciation (*either*, *almond*, *zee/zed*, etc.), vocabulary (*tap/faucet*, *napkin/serviette*, etc.), grammatical usage (*dived/dove*, *sneaked/snuck*, etc.), and spelling (*grey/gray*, *centre/center*, etc.) questions. Similarly, many questions from Chambers' (1994, 1995) Dialect Topography project were included, such as

dinner/supper and the pronunciation of *asphalt*. Several lexical items also came from NARVS (Boberg 2005), including *bathroom/washroom* and *soda/pop*. I also drew from Avis's (1954, 1955, 1956) lexical, syntactic, and pronunciation studies of the Ontario-U.S. border (*progress*, *produce*, etc.), as well as Boberg's (2000) foreign (a) variables (*taco*, *drama*, *avocado*, etc.). Novel questions targeted expected Canadian-American differences, such as the pronunciation of *niche*, the use of *parentheses/brackets*, and the construction *I'm done (with) my homework*. The full list of language questions and variant choices is included in Appendix A.

For most questions, a list of variants was given and a single response was allowed; in two instances (Q16: *waterway*; Q17: *pardon?*), multiple selections were allowed. For a handful of the vocabulary questions, a fill-in-the-blank method was used instead. These decisions were made primarily to keep questions in line with their original counterparts, particularly for variables derived from the Chambers (1994) questionnaire which varies in its question format. It also serves as an attempt to balance the benefit of limited, prespecified choices with the benefit of unbiased free response. Each vocabulary question with answer options also had an "other" option where participants could write their own answers. Spelling questions similarly allowed participants to answer "either one," in line with Scargill & Warkentyne's (1972) version of these questions. The pronunciation and grammar variables did not easily lend themselves to similar options and were therefore limited to single prespecified answer choices. The format of each individual question can be seen in Appendix A.

Participants were also asked for basic demographic information: region, hometown, current town, age, gender, and education. They also provided information on how often they crossed the border for the purposes of work, visiting family/friends, and leisure. At the end of the questionnaire, participants were given three sets of Likert scales (1-5) about personal ideologies. The first set asked participants to rate the similarity of Canadians and Americans, Canadian and American culture, and Canadian and American accents. The second set asked them to rate how pleasantly they perceived Canada/the United States, Canadian/American people, and the way Canadians/Americans speak. The final set asked them to rate how positively they feel about their hometown, moving to another city in their own country, and moving across the border. These scales were based on Swan's (2020) study of the Seattle-Vancouver border which included many measures of attitude and local affiliation; while her study involved in-depth sociolinguistic interviews, the Likert scales are a simpler way to elicit attitudinal information in a questionnaire.

2.2.3 Data analysis

To make aggregate analyses across variables and categories possible, the responses were recoded. As the focus is on the major speech differences separating Canadian and American English, I coded each variant as "Canadian," "American," or "Other." This allowed me to compare the frequency of American and Canadian usage by region across different variables. These codes were based on previous findings from the original studies, as well as dictionary records of American and Canadian or British usage when necessary. A small selection of variables was excluded from analysis due to issues with clearly coding variants; these are noted in Appendix A. Since these variables did not clearly fit the Canadian-American distinction, their exclusion should not affect the analysis of these differences.

In some cases, these divisions have become outdated with certain variables shifting towards American-dominant usage more generally in Canada; however, all participants had the same coding criteria and thus will be affected equally by such variables. These cases may additionally reveal age effects. For example, the majority of Canadians and Americans in this study use the American pronunciation of *leisure* /liːʒəɹ/ rather than the British pronunciation /lɛʒəɹ/. This was also the case in Scargill & Warkentyne's (1972) survey but the proportion of /lɛʒəɹ/ was much greater and showed an apparent time increase in younger speakers. Thus, maintaining traditional Canadian/British coding serves as a reference for the current trajectory of the variants.

Each participant's Canadian and American usage scores were then calculated as the sum of Canadian and American variants, respectively. These scores were calculated for the total dataset as well as each linguistic category (pronunciation, vocabulary, grammar, spelling). "Other" variants were not included in the scores, though they do affect scores by decreasing Canadian and/or American usage for speakers with high frequencies of "Other."

Following Nerbonne (2009), I first focused on aggregated data to identify general trends. This involved comparing mean Canadian and American usage scores between the three regions, first for the whole dataset and subsequently for each linguistic category. For each of these sets of data, ANOVA tests were conducted to determine if region affected Canadian or American usage scores. In all cases, the variance between regions was significant ($p<0.001$). Post-hoc Tukey tests were then performed to see the differences between each pair of regions: Niagara, ON – Niagara, NY; Niagara, ON – GTA; and GTA – Niagara, NY. The full ANOVA and Tukey test results are presented in Appendix E. I also looked at differences for individual variables using Fisher's exact

test. This test was used rather than a traditional chi-squared test as it can handle responses with fewer than five occurrences, which were expected for at least some Canadian variants which Americans rarely use. These findings characterize the main research questions and will constitute the bulk of the results section, which will look at overall and category-specific regional comparisons (2.3.2–2.3.6).

Gender differences (ANOVAs and Tukey tests) and correlations for year of birth (Pearson's *r*) will be discussed alongside the main results. For the ideological data, each scale's Pearson correlation with the various language measures was then calculated for the three regions. The ratings themselves were also analyzed to see, for example, how positively Canadians viewed Americans and their speech on average. The ideological results are discussed in section 2.3.7.

2.3 Results

2.3.1 Demographics

In total, 909 responses were received where at least one language question was answered. Of these, 45 were removed for not meeting selection criteria such as growing up in one of the three regions or being a native English speaker. A further 61 respondents were set aside for currently living in a different region of study than their childhood region. Therefore, 803 responses were available for analysis: 422 (52.5%) from Niagara, ON; 289 (36%) from Niagara, NY; and 92 (11.5%) from the Greater Toronto Area. Though the Toronto group was much smaller, this region is not the main focus and the data is still sufficient for comparison. The distribution of participants is summarized in Table 1, broken down by region and gender.

Region	**Niagara, ON**					**Niagara, NY**			**Greater Toronto Area**	
Gender	**F**	**M**	**NB**	**O**	**N/A**	**F**	**M**	**NB**	**F**	**M**
Participants	330	84	5	1	2	211	76	2	67	25
Total	422 (52.5%)					289 (36%)			92 (11.5%)	

Table 1: Summary of participants. F = female; M = male, NB = non-binary, O = other gender, N/A = no info.

The participants are moderately skewed in terms of education and gender. There is an unusually large number of participants with or in the progress of obtaining graduate degrees (26.2%), though the high school (14.8%), vocational/technical school (25.7%), and bachelor's degree (32.9%) groups were relatively well-represented. The gender of participants was

overwhelmingly female (75.7%). There were however many males (23%) as well as some non-binary (0.9%) and unlisted gender (0.1%) respondents[1]. This gender gap is not unexpected given the uncontrolled nature of an online survey.

The overall year of birth ranged from 1932 to 2004, with a mean of 1970 (approx. 52 years old) and median of 1969 (approx. 53 years old). The sample is therefore slightly on the older side, but a large spread of ages is represented. In Niagara, ON, the year of birth distribution is similar to the overall dataset (range: 1941–2003, mean: 1972, median: 1971). The Niagara, NY participants tended to be older (range: 1932–2004, mean: 1962, median: 1960) while the GTA participants tended to be younger (range: 1945–2003, mean: 1983, median: 1986). Despite some overrepresented categories, the distribution of participants is sufficient for the regional, age, and gender analyses I will focus on here.

2.3.2 Overall regional comparison

An aggregate analysis can reveal whether there is any diffusion pattern overall: are the Canadians always different from the Americans or is their usage ever the same? The effect of region on aggregate Canadian and American scores was analyzed across the entire dataset as well as each linguistic category subset. As mentioned, all ANOVAs for region were significant for these aggregate scores ($p<0.001$ for each). The main analysis will therefore depend on Tukey tests which show whether each pair of regions is significantly different from each other. Results are reported in Table 2 for overall, vocabulary, pronunciation, grammatical, and spelling aggregates. The means and standard deviations for Canadian and American scores in each region are reported in Table 3.

Variables		**Linguistic Category**				
Regions	**Score**	**Overall**	**Vocabulary**	**Pronunciation**	**Grammar**	**Spelling**
NON – NNY	Can.	**	**	**	**	**
	Am.	**	**	**	**	**
NON – GTA	Can.	-	-	*	-	-
	Am.	-	-	*	-	-
GTA – NNY	Can.	**	**	**	**	**
	Am.	**	**	**	**	**

Table 2: Results for post-hoc Tukey tests between each pair of regions for overall, vocabulary, pronunciation, grammar, and spelling for Canadian (Can.) and American (Am.) usage scores. * = $p<0.05$; ** = $p<0.001$. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

[1] No education data was reported for 0.4% and no gender data was reported for 0.2%.

Variables		Linguistic Category									
		Overall		Vocab.		Pron.		Grammar		Spelling	
Regions	Score	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NON	Can.	18.7	3.4	8.7	1.9	6.2	2.2	2.0	0.8	2.8	0.8
	Am.	20.9	3.4	6.2	1.9	9.7	2.3	2.9	0.9	1.2	0.8
GTA	Can.	19.6	3.3	8.9	1.9	6.8	2.2	2.0	0.9	2.7	0.8
	Am.	20.2	3.7	6.2	1.9	9	2.3	2.8	0.9	1.3	0.8
NNY	Can.	8.6	2.8	1.8	1.4	2.6	1.9	1.1	0.8	1.4	0.7
	Am.	30.9	3.7	13.2	2.0	13	2.5	3.8	1.0	2.7	0.8

Table 3: Mean and standard deviation (SD) of Canadian and American usage scores in each region for overall (max=43), vocabulary (max=17), pronunciation (max=16), grammar (max=5), and spelling (max=5). SD = standard deviation (rounded to one decimal place). GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

The first observation from these results is that, in all cases, Niagara, NY is significantly different from both Canadian regions. The direction of these differences indicates that New York is always more American and less Canadian. This is expected since the variables are intended to show Canadian and American differences. Importantly, this indicates that Niagara, ON is never as American as Niagara, NY since they always differ. Though Canadians use slightly more American than Canadian variants overall, they do not use nearly as many as the Niagara, NY group. Therefore, the border clearly divides American English from Canadian English overall.

When looking at these overall aggregate scores, there is an interesting correlation with year of birth: across all participants, there is a small positive correlation with the Canadian scores aggregate (r=0.3, p<0.001) and a small negative correlation with the American scores aggregate (r=-0.27, p<0.001). This indicates that, overall, younger participants use more Canadian and fewer American variants than older participants. However, no region on its own shows either correlation, suggesting some general change in usage across the regions that is not particularly strong. No gender difference was found.

The more interesting outcomes are the differences between Niagara, ON and the GTA. The overall aggregate comparison shows that the Canadian regions generally have the same frequency of Canadian and American variant use. It is also immediately apparent that most of the individual linguistic categories do not show significant variation between Niagara, ON and the GTA. However, the difference in pronunciation aggregate scores is significant for both American and Canadian usage. In this case, Niagara speakers use fewer Canadian variants and more American variants. In both regions, the pronunciation means clearly favour American variants with Niagara,

ON ahead of the GTA. However, the Canadians do not use American pronunciations as frequently as Americans do since the differences between both Canadian regions and Niagara, NY are also significant. This divide between pronunciation and the other linguistic categories reveals a difference in which types of features are diffused across the border: while American pronunciations may be adopted in place of traditional Canadian pronunciations, the Canadian vocabulary and spelling differences appear to be generally maintained. Grammar shows some preference for "American" variants, though as we will see these variables are largely outdated.

In all cases, the aggregate only shows the general pattern. In reality, some variants are strongly maintained while others have become overwhelmingly American, and thus contribute differently to the overall usage in their respective categories. To find the actual cases of diffusion, we must look at each linguistic category separately and analyze key individual variables. The next four sections will delve into the specific patterns of diffusion and non-diffusion for pronunciation (2.3.3), vocabulary (2.3.4), spelling (2.3.5), and grammar (2.3.6). I will focus here on the key variables; the response frequencies for all variables can be found in Appendix A.

2.3.3 Pronunciation

The pronunciation variables are the only ones for which the border does not overall act as a strict barrier. Though the Niagara, ON group still differs from the Americans, the fact that they overall use more American and less Canadian pronunciations than the GTA suggests there is at least some wave diffusion across the border. It should be noted that the pronunciation variables studied in the questionnaire are not related to phonemic inventory differences and therefore do not require speakers to undergo structural changes to their linguistic systems. Instead, these differences are lexically-based and can be perceived and learned relatively easily.

To look more closely at the diffusion of American pronunciation, each variable can be examined individually. This will reveal the most important variants that contribute to Niagara's higher American and lower Canadian usage scores. Each variable was therefore compared between each pair of regions. A simplified breakdown of regional differences is presented in Table 4.

	Regions		
Variable	**NON-NNY**	**NON-GTA**	**GTA-NNY**
Zee/Zed	**	-	**
Leisure	-	-	-
Asphalt	**	*	**

Either	**	-	*
Produce	**	-	**
Project	**	-	-
Progress	**	-	**
Again	**	-	**
Almond	*	-	-
Caramel	**	*	**
Niche	**	*	**
Pasta	**	**	**
Drama	**	-	*
Taco	*	-	*
Llama	**	-	**
Avocado	*	-	*

Table 4: Simplified *p*-value results for Fisher's tests between each pair of regions for each pronunciation variable. * = $p<0.05$, ** = $p<0.001$. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

There are a few striking observations from the individual variable results. First, there are only a few instances where Niagara, NY does not significantly differ from the Canadian regions. This confirms that, in general, no Canadian region uses enough American pronunciations to be considered "just as American" as New York. There is only one variable (*leisure*) where Niagara, ON does not differ from Niagara, NY; however, the GTA speakers also do not differ from Niagara, NY. This suggests that the British pronunciation /lɛʒə(ɪ)/ is simply not a strong marker of Canadian English anymore, at least in the regions studied. Beyond this, we also see that the GTA actually has more similarities (non-significant differences) with Niagara, NY than Niagara, ON does. In addition to *leisure*, the GTA also uses more American pronunciations for *project* and *almond*. These variables thus show hierarchical diffusion, with Toronto adopting the American pronunciations more quickly than Niagara, ON. This contrasts with the aggregate pronunciation finding that Toronto is generally more Canadian and less American than Niagara, showing that not all variables follow this trend.

Additionally, the variables which show significant differences in usage between Toronto and Niagara (*asphalt*, *caramel*, *niche*, and *pasta*) differ in their patterns of diffusion. To see the direction and degree of these differences, the frequency of each variant in each region is shown in Tables 5-8 with the coding in brackets (A: American, C: Canadian). Niagara, NY is included for ease of comparison. Participants who did not provide a response are excluded.

Table 5: In ASPHALT, does the AS sound like **ash**?

Region	Yes (C)	No (A)
Niagara, ON	88.7%	11.3%
Greater Toronto Area	79.1%	20.9%
Niagara, NY	37.6%	62.4%

Table 6: Is CARAMEL pronounced with two syllables or three syllables?[2]

Region	Two (A)	Three (C)
Niagara, ON	26.5%	73.5%
Greater Toronto Area	10%	90%
Niagara, NY	55.3%	44.7%

Table 7: Does NICHE rhyme with **ditch** or **quiche**?[3]

Region	Ditch (A)	Quiche (C)
Niagara, ON	8.6%	91.4%
Greater Toronto Area	1%	99%
Niagara, NY	43.6%	56.4%

Table 8: Does the first part of PASTA sound like the first part of passing or possible?

Region	Passing (C)	Possible (A)
Niagara, ON	42%	58%
Greater Toronto Area	72.5%	27.5%
Niagara, NY	12.4%	87.6%

From these individual breakdowns, it becomes clear that not all variables involve greater American usage in Niagara, ON than in Toronto. For *asphalt*, the GTA group uses more of the American variant than the Niagara group. These variables may therefore show more hierarchical diffusion in progress. However, the wave pattern appears for the other variables: Niagara, ON has greater frequency of two-syllable *caramel*, *niche* rhyming with "ditch", and the /ɑ/ pronunciation

[2] *Caramel* is evidently split in Niagara, NY with a slight preference for two syllables. The coding in this case assumes three syllables is generally used in Canada but two is not, making the three-syllable version the preferred "Canadian" variant; it is not meant to suggest it is solely a Canadian variant.

[3] *Niche* is also quite variable in Niagara, NY. While the "ditch" pronunciation is typically found in the U.S. but not Canada, "quiche" also occurs in the U.S. The Merriam-Webster dictionary (Merriam-Webster 2023) lists "ditch" first before "quiche" and "dish"; the Cambridge dictionary (Cambridge University Press 2023) lists "ditch" as the American pronunciation.

for *pasta*[4], all in line with American usage. The starkest contrast is *pasta* (p<0.001) where the majority response is actually different. Niagara's frequency of the American "possible" variant is more than double the frequency for the GTA. Though the frequency in Niagara, ON is still not very close to that in Niagara, NY, this variable seems to be a clear example of wave diffusion across the border. The same is true of *caramel* and *niche*, though to a lesser extent. Therefore, Niagara, ON seems to act as an intermediate dialect area for these variables.

Pasta is the only foreign (a) variable which showed a difference, but this is actually because the others (*drama, llama, taco, avocado*) all have very high use of the American /ɑ/ forms in Canada. These are the majority variants in both regions, ranging from 73-94% of participants (see Appendix A), though they still differ significantly from the Niagara, NY usage (as seen in Table 4). The American foreign (a) variants therefore seem to be rapidly diffusing, as reported by Boberg (2000), but *pasta* is the only one where the actual pattern of diffusion is clear. I will return to these variables and more foreign (a) words in the acoustic study (Chapter 3).

There are also several other pronunciation variables where the Canadians use a majority of the American variant: *either* (like "bee"), *progress* (like "got"), *produce* (like "go"), *project* (like "got"), *again* (like "pen"), and *almond* (with /l/) (see Appendix A). These variables may therefore show diffusion with no clear wave or hierarchical pattern. Like the foreign (a) variables, these all still significantly differ from Niagara, NY usage.

As with the overall scores, no gender differences were found for the pronunciation aggregate scores across or within regions. The apparent time results however are striking. In all three regions, the use of American pronunciation increases with year of birth. In both Canadian regions, Canadian pronunciation also decreases. This is contrary to the overall pattern in which Canadian usage increases and American usage decreases, though that correlation only arose when all regions were considered together. The pronunciation correlations are reported in Table 9.

	American Usage	**Canadian Usage**
Niagara, ON	r = 0.37 **	r = -0.42 **
Niagara, NY	r = 0.12 *	r = -0.11, p=0.067
Greater Toronto Area	r = 0.39 **	r = -0.55 **

Table 9: Pearson correlation results between American and Canadian usage scores and year of birth in each region for all pronunciation variables. Positive values indicate an increase in apparent time. * = p<0.05, ** = p<0.001.

[4] It should be noted that the exact quality cannot be determined in a written questionnaire and may not be the same as the American pronunciation (see Boberg, 2009). I return to this issue in the acoustic study.

These results are in line with the fact that the pronunciation variables differ from the other categories, given the aggregate regional differences and the diffusion of several variables.

2.3.4 Vocabulary

In contrast with pronunciation, vocabulary shows little diffusion. Figure 1 shows the mean American and Canadian vocabulary usage scores in each region. While there is moderate use of American variants by Canadians, and certainly more than Canadian variant use by Americans, there is still a clear preference for Canadian terms.

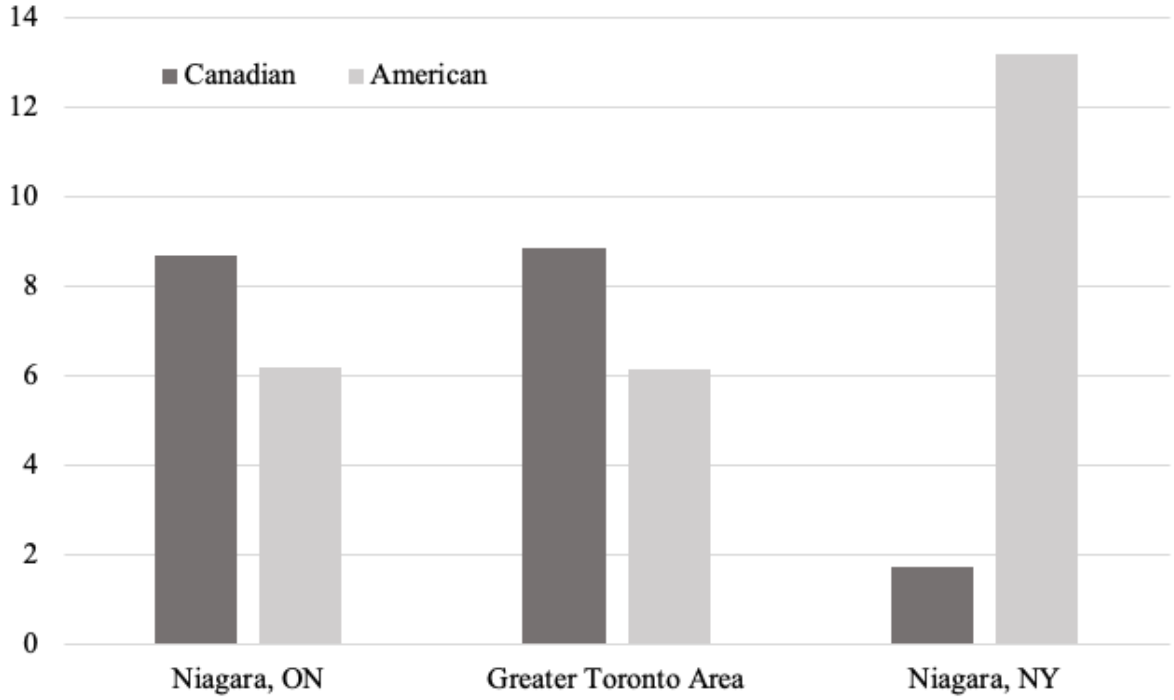

Figure 1: Mean American and Canadian vocabulary usage scores by region. Max = 17 coded variables.

For a couple of variables (*napkin*, *dinner*[5]), Canadians do show majority use of the American term. However, the vast majority of vocabulary variables favour Canadian usage. The strongest differences, summarized in Tables 10-14, show a sharp split. The use of *brackets* rather than *parentheses* in Canada (Table 14) is of particular interest since it has not been included in previous Canadian English surveys; the results reported here show a clear divide. Note that only the top variants are included in each table (i.e., not all values add up to 100%).

[5] This variable is not strictly Canadian vs. American. Jankowski & Tagliamonte (2019) show that *dinner* use is generally higher in Toronto and Loyalist settlements like Niagara, suggesting it is already more common in the Canadian regions studied here.

Table 10: When you want water from the sink, you turn the _____:

Region	Tap (C)	Faucet (A)
Niagara, ON	84.4%	11.1%
Greater Toronto Area	87%	10.9%
Niagara, NY	14.6%	74.3%

Table 11: What do you call the thick, sweet spread used to cover a cake?

Region	Icing (C)	Frosting (A)
Niagara, ON	81.9%	16.9%
Greater Toronto Area	76.1%	20.7%
Niagara, NY	18.4%	80.6%

Table 12: What do you call the metal structure around the edge of the roof that collects rainwater?

Region	Eavestroughs (C)	Gutters (A)
Niagara, ON	86.5%	12.6%
Greater Toronto Area	73.9%	25%
Niagara, NY	3.8%	95.5%

Table 13: When you've finished your meal at a restaurant, you ask for the _____:[6]

Region	Bill (C)	Check (A)	Cheque
Niagara, ON	62.5%	15.9%	19.2%
Greater Toronto Area	71.7%	5.4%	21.7%
Niagara, NY	16%	80.6%	1%

Table 14: What do you call the following symbol: **( )** ?

Region	Brackets (C)	Parentheses (A)
Niagara, ON	71%	24.6%
Greater Toronto Area	62.6%	29.7%
Niagara, NY	1.8%	96.3%

Turning to the social variables, we do see some gender and year of birth effects for vocabulary use. The only gender difference is in Niagara, ON where women have higher Canadian usage than men, though the difference is relatively small (p<0.05; female mean: 8.8, male mean:

[6] Note that while the majority of Canadians in both regions do use the Canadian term "bill," a sizeable number use the American term "check" but prefer the Canadian spelling "cheque."

8.2). The apparent time results are more interesting: across all speakers, there is an increase in Canadian usage (r=0.27, p<0.001) and decrease in American usage (r=-0.26, p<0.001). This would suggest higher use of Canadian terms and lower use of American terms among younger speakers. However, the opposite pattern emerges when looking at the three regions individually. In Niagara, ON, there is a very small increase in American usage (r=0.1, p<0.05) and decrease in Canadian usage (r=-0.1, p<0.05). In the GTA, only the decrease in Canadian usage is significant (r=-0.28, p<0.01) but it is slightly stronger. Neither pattern is significant in Niagara, NY. This raises the question of why there is an overall increase in Canadian usage if there is a decrease in two of the individual regions. This may be explained by the differences in year of birth between regions: since the Niagara, NY participants use fewer Canadian variants overall *and* tend to be older, the combined correlation will be skewed (e.g., more older speakers have lower Canadian scores due to being American, not due to age). The correlations by individual region are therefore more reflective of the apparent time data, suggesting a potential small decrease in Canadian vocabulary usage in both Canadian regions and no change in New York.

2.3.5 Spelling

The spelling variables were the most clearly split: every Canadian variant was preferred by the majority of Canadians. This division was particularly strong for *colour* vs. *color*, *cheque* vs. *check*, and *travelled* vs. *traveled*, with preference for the national variant ranging from 74%-94% (see Appendix A). The less clear variables were *centre* vs. *center* and *grey* vs. *gray*. For *centre*, only 54% in Niagara, ON and 63% in the GTA chose the Canadian variant *centre*, with 30% in both regions instead choosing "either one"; on the other hand, Niagara, NY clearly preferred *center* at 91%. For *grey*, the Canadians responded similarly with 52-54% *grey* and 35-38% "either one". In this case however, the Americans had no preference and in fact the most popular response (49%) was "either one." This may simply indicate that *grey* vs. *gray* is not a strong marker for Canadians or Americans.

There are no gender differences for spelling use. For year of birth, we again see some conflicting correlations. Across all speakers, American usage slightly decreases in apparent time (r=-0.23, p<0.001) and Canadian usage slightly increases (r=0.17, p<0.001). This seems to suggest that younger speakers are using more Canadian spelling and less American spelling. However, the only significant correlation in an individual region shows a slight decrease in Canadian usage in

Niagara, ON (r=-0.14, p<0.01). There may therefore be some wavelike diffusion of American variants in progress in Niagara, ON, though the frequencies still remain very distinct.

2.3.6 Grammar

The grammar section is essentially the opposite of spelling: speakers use the same majority variant in all three regions for four of the five variables: *snuck* over *sneaked*, *different than* over *different from*, *lent* over *loaned*, and *dove* over *dived*. The proportion of participants using the majority term is also within 1-4% between each region, with the exception of *lent* where the GTA usage is lower (NON: 75%, GTA: 54%, NNY: 77%; see Appendix A). The overall border difference then seems surprising: if most of the variables are very close in usage across regions, why are the Canadians overall different from the Americans? The difference is likely driven by the fifth variable which has a large divide, with Canadians preferring *done* over *done with* (Table 15). Since there are only five variables in the grammar category, one variable makes a greater difference in the aggregate scores than individual variables do in the other categories.

Table 15: Which do you say?

A. I'm done my homework.

B. I'm done with my homework.

Region	done (C)	done with (A)
Niagara, ON	81%	19%
Greater Toronto Area	78.7%	21.3%
Niagara, NY	2.9%	97.1%

We may then expect the aggregate regional difference to disappear if the *done (with)* variable is removed. This is true: there is no significant regional difference in that case, though it remains close (F(2)=2.768, p=0.0634). These findings are in line with older questionnaires which showed Canadians leaning towards the "American" forms rather than those associated with British or Old English (Avis, 1955; Scargill & Warkentyne, 1972). Though the older variants are coded here as "Canadian" to examine the aggregate differences, they do not have the strong Canadian association that most of the pronunciation, vocabulary, and spelling variables have.

It is therefore not surprising that grammar shows no correlation with year of birth, whether or not the *done (with)* variable is removed; this category is largely stable and has been for many years. There is a gender difference in Niagara, ON: women have significantly higher Canadian usage than men (p<0.01, whether or not *done (with)* is removed) and lower American usage

(p<0.05, only if *done (with)* is removed). However, since there are only five variables any difference is quite small and may not be very meaningful.

2.3.7 Likert scales

Pearson correlation coefficients were calculated between the three social scale sets (positivity towards Canada, positivity towards America, sense of affiliation to hometown/country) and the language scores. Contrary to the hypothesis, most of the ideological scales were not correlated with Canadian or American usage scores. When looking at all participants, only affiliation was weakly associated with language use (American scores: r=-0.13, p<0.05; Canadian scores: r=0.15, p<0.01). These correlations indicate that speakers with a greater sense of affiliation to their country/hometown use fewer American forms and more Canadian forms. This finding is similar to Swan's (2020) results which suggested that ties to one's city and opposition to newcomers were linked to greater use of the local variant. However, since the majority of responses are from Canadians, affiliation was investigated by region. In this case, no significant correlations emerged suggesting that the overall correlations may not be very meaningful.

Considering the patterns of diffusion were mostly linked to pronunciation, those variables were also investigated independently. None of the scales were significantly correlated when all regions were considered together. However, some weak correlations were found for Niagara, ON on its own: American usage scores were positively correlated with positive views about both Canada (r=0.15, p<0.01) and America (r=0.13, p<0.01). Canadian scores were also negatively correlated with positive views about both Canada (r=-0.11, p<0.05) and America (r=-0.12, p<0.05). Combined, these seem to indicate that Niagara, ON speakers who view America and Americans more positively use slightly more American forms and fewer Canadian forms. However, those with positive views of Canada also follow this pattern. These correlations are quite weak but suggest that attitudes towards each country may be related to variant use and provide a basis for further study. No other regions had significant correlations for pronunciation with any ideological scale.

I also examined the results for the scales about how pleasant Americans and Canadians "sound" to see general attitudes towards American and Canadian English. Canadians ranked American speech as significantly less pleasant (p<0.05) than Canadian speech. Further, only 28% of the Canadians gave American speech a positive score (4-5), compared to 79% for Canadian speech. Canadians therefore seem to generally find American English less pleasant.

2.4 Analysis

The overall patterns in the questionnaire data show that the border primarily acts as a barrier to diffusion. In all cases, the Canadian regions remain significantly different from New York at the aggregate level. Though some individual variables favour the American variant, only a few of these are just as frequent in Canada as they are in Niagara, NY (e.g., *leisure* or the traditional grammar questions). Considering the proximity of Niagara, ON and the size of Toronto, neither are as similar to Niagara, NY as they should be if diffusion is simply a function of the distance and size of communities. We may then conclude that the border is largely a barrier to diffusion, or at a minimum that it reduces the degree of diffusion.

However, the pronunciation data do show some diffusion-like trends across the border. Many of the variables have majority use of the American variant in all three regions. Compared to older studies, many of these have apparently increased in usage among Canadians. *Leisure* rhyming with "pleasure" was already falling out of favour in Canada in Scargill & Warkentyne's (1972) survey, but their Ontario respondents chose this variant around 20% of the time while the current participants chose "pleasure" 7% (NON) and 12% (GTA) of the time. Similarly, *again* rhyming with "pane" was declining in their survey at around 50% for the younger Ontario participants; this has fallen to 24% (NON) and 21% (GTA) in my survey. The pronunciation of *produce* with the vowel of "got" has also drastically declined compared to Avis's (1956) study, from 71% to 19% (NON) and 30% (GTA). However, *either* with the vowel of "pie" shows very similar rates (around 30%) to those in Scargill & Warkentyne's (1972) report, contrary to their belief that "[*i* of *bide*] will edge out [*ee* of *beet*] completely in a few generations" (63). *Progress* with the vowel of "go" has also stayed relatively stable (around 40% in both surveys).

The foreign (a) pronunciation variables also generally show diffusion, with all except *pasta* greatly favouring the American /ɑ/ in both Canadian regions. For *pasta*, we see that the Canadian speakers closer to the border (Niagara, ON) use more of the American pronunciation than Canadians further from the border (GTA). These findings are in line with Boberg's (2000) results for foreign (a) pronunciation in Windsor versus Toronto. This pattern corresponds to the wave diffusion model, as the American pronunciation spreads gradually across geographical space rather than spreading hierarchically to major cities like Toronto first. It is possible that the other foreign (a) variables also diffused in a wavelike pattern but that they are now too far along to see a difference between Niagara, ON and the GTA. The five variables chosen for the questionnaire

provide only a limited sample of this phenomenon; the foreign (a) list is expanded in the acoustic study to see if Niagara, ON is more generally further ahead.

These pronunciation variables are also particularly interesting since the correlations show an apparent time increase in American usage and decrease in Canadian usage in both Niagara, ON and the GTA. This is contrary to the overall aggregate pattern and these correlations are quite a bit stronger at around r=0.4-0.5 in both Canadian regions. Younger speakers therefore chose more American variants overall. This seems to be in line with the variables discussed above which mostly show increases in American usage compared to previous surveys. While the Canadians are still more different from the Americans than we might expect without the border, they are gradually adopting American pronunciations.

However, not all of the variables showed majority American usage. *Asphalt*, *caramel*, and *niche* all favour the Canadian variant despite their differences between Niagara, ON and the GTA. Toronto also continues to favour /æ/ in *pasta* despite the high rates of /ɑ/ in the other foreign (a) variables. Additionally, *zed* continues to be highly favoured over *zee* in both Canadian regions (73% in Niagara, ON; 78% in the GTA), in contrast with Burnett's (2006) findings where *zee* was the majority variant in the border town of St. Stephen, New Brunswick. Thus, while there are some patterns of diffusion of American pronunciations, this depends largely on the variable. Returning to the main research question, this suggests that Niagara, ON acts as an intermediate area in a dialect continuum for pronunciation at the aggregate level (wave diffusion), but that individual variables may show wave, hierarchical, or no diffusion. We must also recall that Niagara, NY differs significantly for overall pronunciation and for the vast majority of variables, even if Canadians also favoured the American variant, which would not be expected if the border had no effect on diffusion. As such, the variables that do show diffusion are not typical wave or hierarchical diffusion; they are diminished by the border.

The grammatical usage questions also emphasize the importance of the variable being studied. As expected, most participants use the "American" form, in line with previous findings (e.g., Avis 1955, Scargill & Warkentyne 1972), and the regional differences disappear completely without the *done (with)* variable. This category is more difficult to code as "Canadian" or "American" as the non-majority variants tend to be seen as non-standard and their use also depends on education and rurality (Scargill & Warkentyne 1972). These variables are therefore not as strong of an indicator of diffusion as pronunciation. Additionally, use of *done* rather than *done*

with does show a strong Canadian-American contrast. This variable was not included in the previous questionnaires so it is not possible to compare, but the large difference in usage suggests that the border is a barrier in this case. To actually see patterns of diffusion in grammatical usage, we would need to examine more variables that do still show a Canadian-American contrast and see if usage differs in Niagara; in this case, it does not.

The border also acts as a barrier overall for vocabulary and spelling. In both cases, the majority of variables show a strong American-Canadian contrast. Though the spelling correlation in Niagara, ON may suggest some level of wave diffusion as American variants show a small increase there, the fact that all variables show preference for the Canadian form greatly favours the border acting as a barrier for spelling. Similarly, vocabulary shows a small potential decrease in Canadian usage in both Canadian regions, but most variables continue to favour the Canadian variant. Combining *done (with)* and the several pronunciation variables that favour Canadian usage with the vocabulary and spelling findings, we see that Canadian forms are generally still preferred in Canada. These findings are in line with the previous studies outlined in section 1.2.1: though certain variables that tend to diffuse are identified, this is typically in contrast with a majority of variables that do not.

While these results show a fairly clear current pattern, the apparent time results are overall unclear. With the exception of pronunciation, discussed above, the correlations are weak and often conflicting. Across all regions, there is a general tendency for younger speakers to use more Canadian and fewer American forms. However, many of the individual region correlations show the opposite tendency. This is not clearly driven by any particular region though it may be a result of the different age distributions in each region. It is therefore difficult to make any strong conclusion about apparent time change; some cases show potential decrease of Canadian forms while others show potential increase. Given the large divide in many variables and the overall aggregate showing no regional correlations, it seems there is little change related to year of birth outside of pronunciation.

The maintenance of many Canadian forms could be related to maintenance of local identity, as has been suggested in previous studies (e.g., Miller 1989, Burnett 2016, Swan 2020). This is supported by the negative ratings Canadians tended to give American speech in comparison to their positive ratings of Canadian speech. Of course, these scales are not particularly informative

as we do not know what aspects the participants are rating speech on. I investigate this question further in the acoustic study interviews.

In general, the fact that the results include wave diffusion, hierarchical diffusion, and non-diffusion makes it clear that all three need to be considered in the study of diffusion. Additionally, the border has an effect even in the cases of diffusion since the Canadians still tend to remain significantly different from the Americans. The actual effect of the border also depends on the variable being studied: I have shown different patterns for pronunciation and grammar in comparison with vocabulary and spelling, as well as different patterns for individual variables within each category. In the next chapter, the acoustic data will further show that the effect of the border varies. In order to accurately predict diffusion, these varying patterns and the general weakening effect the border has on diffusion must be taken into account.

3. Acoustic study

3.1 Introduction

In the acoustic study, I interviewed speakers from the three regions to further investigate the diffusion and non-diffusion of pronunciation across the border. In this case, phonetic variables which cannot be studied via questionnaire were the primary target. Foreign (a) variables were also included to see if the results of the questionnaire accurately reflected the pronunciation used by participants. Participants were also asked more in-depth questions about their attitudes towards Canadians, Americans, and Canadian and American English.

Though the whole vowel space was measured for each participant, I will focus on the vowels which are most relevant to a Canadian-American comparison as outlined in section 1.2.2. The first comparison deals with the Northern Cities Shift. Since I am focusing on Canadian patterns, I will not do a detailed analysis of this shift. Instead, I will look at the basic components to see the pattern in each region. One important component is whether TRAP (/æ/) is tensed and thus raised and fronted; since only the nucleus is measured, we will not see the inglide associated with the NCS short-*a* tensing (Labov et al. 2006) but will still see how raised and fronted /æ/ is relative to the Canadian vowel. The second is whether LOT-THOUGHT (/ɑ/-/ɔ/) is unmerged and whether unmerged LOT (/ɑ/) is fronted.

The low back (LOT-THOUGHT) merger will also form the basis of another comparison, namely the Low-Back-Merger Shift. This shift of the short front vowels is thought to be triggered by a gap in the low central vowel space after LOT merges at the back with THOUGHT (Labov et al. 2006). Assuming the Canadians retain the merger, they are expected to show the shift with retraction of KIT (/ɪ/), DRESS (/ɛ/), and TRAP (/æ/); the Americans are expected to have no merger and thus no shift. However, since the Northern Cities Shift also involves retraction/lowering of /ɪ/ and /ɛ/, this comparison will depend on the retraction of /æ/ as well as the merging of /ɑ/-/ɔ/.

The next comparison looks at the raising of allophones of /æ/ in pre-velar (/æG/) and pre-nasal (/æN/) environments. The distance between /æ/ and these allophones is expected to be much greater for Canadians, with a low and retracted /æ/ in comparison to raised and fronted /æG/ and /æN/. In Ontario specifically, we expect /æN/ to be further raised than /æG/ (Boberg 2008, 2010). For the Americans, we expect all allophones of /æ/ to be tensed and thus raised and fronted as a group, though /æN/ may be higher/fronter within this group (Labov et al. 2006).

I will then look at the Index of Phonetic Innovation which uses the relative fronting of /u:/ and retraction of /æ/ as a measure of how innovative or conservative speakers are. Since the Americans are expected to have tensed /æ/ and the Canadians are expected to have retracted /æ/, this measure should show Canadians to be more innovative than Americans. A comparison between the Canadian regions will also show whether Niagara, ON and GTA speakers are equally innovative.

Canadian Raising (/au/ and /ai/ before voiceless obstruents) will then be analyzed. I will look at whether the Canadians and Americans raise /ai/ and/or /au/, comparing the overall degree and dimensions (i.e., height and fronting) of raising. As noted in 1.2.2, Americans in Upstate New York have previously shown /ai/-raising but do not typically show /au/-raising, while Canadians are expected to show both. Additionally, I noted many apparent instances of raising for *houses* and *mouths* when interviewing the Canadians despite the following voiced obstruent. These words will be analyzed individually to verify this acoustically.

The final comparison returns to foreign (a), given the results of the questionnaire. I will compare the overall tendency towards front and back (a) pronunciations in each region as well as the actual phonetic quality used. Individual words will also be analyzed to see which ones favour the front or back (a) and whether there are cases of intermediate pronunciation.

In addition to these acoustic comparisons, I will look at the effect of style by comparing the word list and spontaneous speech for a subset of Niagara, ON speakers. I will also summarize the qualitative findings for questions assessing attitudes towards Canadians and Americans and their speech. These questions look at more specific judgments on Canadian and American English, including which features speakers perceive as identifiably American or Canadian.

This chapter will go through the methodology (3.2), results (3.3), and analysis of the results (3.4) for the acoustic study. These findings will then be discussed in conjunction with the questionnaire findings in Chapter 4.

3.2 Methodology

The acoustic study involved acoustic analysis of recorded interviews. These interviews were conducted with participants from each of the three regions to allow for further comparison of the diffusion and non-diffusion of variants across the border. This section will outline the

methodology used for the acoustic study including the participants and data collection (3.2.1), the interview process (3.2.2), and the data analysis methods (3.2.3).

3.2.1 Data collection

As with the questionnaire, participants were recruited from three key regions: Niagara, ON; Niagara, NY; and the GTA. Participants were again native English speakers aged 18 years or older who grew up in one of the three regions. The goal was to interview at least ten men and ten women from each region. There was also a general goal to collect additional data from Niagara, ON as the primary region of study and as an area with little existing data. Participants were primarily recruited from the questionnaire: all questionnaire participants were asked whether they would like to be contacted for the interviews. A recruitment email was then sent to everyone who expressed interest and interviews were scheduled with participants who met criteria. No compensation was offered. Interviews were conducted online using Zoom; I interviewed all participants myself. The meeting was recorded with both combined and speaker-separated audio tracks for analysis.

3.2.2 The interviews

The interview followed the same basic format for each participant. The first part asked participants to read a word list which targeted the vowels of interest and the vowel space more generally. These data are the basis of the main analyses. The words are listed in Appendix B, organized by their target vowel. The second part was a list of sentences. The sentences were presented as questions with the target word at the end to try to ensure the word was accented without drawing attention to it; this did not work as most participants read the questions with flat intonation, and this data will therefore not be analyzed. The third part involved spontaneous speech. Participants were asked about how often they crossed the border and their feelings about living near the border, as well as various questions about work, hobbies, etc. to elicit more speech. The spontaneous speech from a subset of speakers will be used as a comparison with the word list data. Participants were also asked to identify what makes someone sound Canadian or American to them and their general thoughts on how different Canadians and Americans sound.

3.2.3 Data analysis

The acoustic data were transcribed and annotated in Praat (Boersma & Weenink 2021). The speaker-separated audio track of each participant was manually annotated in a TextGrid. Following the methodology from the Atlas of North American English (Labov et al. 2006), I used single-point measurements placed at the maximum F1 for vowels characterized by the fall and rise

of the tongue (e.g., short vowels, long upgliding vowels) or the point of inflection in F2 for vowels characterized by tongue movement to and from the front/rear periphery (e.g., ingliding vowels). Single-point measurements allow for easier comparison between regions across the entire vowel space, whereas trajectories produce a very large number of comparison points for each vowel (Labov et al. 2006). For diphthongs, this also means only the nucleus is measured. While the offglide may provide additional information, the principal difference in the regions I am looking at is in the nucleus – for example, Canadian Raising depends largely on the height of the nucleus in /ai/ and /au/ before voiceless obstruents.

I will deal primarily with the word list data for this analysis. Every token was annotated unless the participant said the wrong word, the word was deaccented, or the recording quality was not sufficient to measure formants. The final word in the list (*wives*) was excluded for all participants due to a tendency to deaccent at the end. For the spontaneous speech that is analyzed here, I also only measured accented words which were sufficient for formant measurement.

The vowels which were annotated in each word from the word list are listed in Appendix B. For *hou(s)es*, the *(s)* was first coded for each participant as /s/ or /z/ using impressionistic and spectrogram analysis, since a subset of speakers (24.6%) used /s/. If the speaker used /s/, the vowel in *houses* was coded as /auT/ rather than /au/ and will therefore not affect the analysis of potential raising before /z/. This alternation seems to be specific to *houses*, though two instances of /s/ in *mouths* were identified; these tokens were excluded.

The F1 and F2 values from each measurement point were then extracted using a Praat script. The formant ceiling was set to 5500Hz for women and 5000Hz for men. For the agender participant, I tested both settings to compare formant tracking and decided to use the 5500Hz setting. I then examined the formant tracking for unexpected values based on the vowel and national identity (following Boberg's (2010) Canadian mean values for Canadians and Boberg's (2021) General American English mean values for Americans as reference). Any unexpected values were manually checked in Praat and formant settings were adjusted if needed to accurately track the visible formants; the values were then manually changed to the corrected measurement.

The vowel formants were then normalized following Labov et al.'s ANAE (2006) method. This method is a speaker-extrinsic modification of the speaker-intrinsic Nearey2 method (Nearey 1978). It also scales the formant values rather than generating non-Hertz values, allowing for a

clear comparison of formants in the entire vowel space. The normalized Hertz will simply be referred to as Hertz (Hz) throughout, but all results use normalized values.

As with the questionnaire, ANOVAs were used to find significant variance between the three regions and post-hoc Tukey tests were then used to look at the significant differences between pairs of regions or between pairs of vowels within each region. This type of analysis looks at F1 and F2 separately, showing the dimension(s) in which the vowels or regions differ across all speakers. Additionally, many of the comparisons use Euclidean distances to measure the combined difference in F1 and F2. This is particularly important for diagonal movement, such as the raising and fronting pattern for Canadian Raising, the raising and fronting of allophones of /æ/, and the retraction and lowering of the short front vowels involved in the LBMS. For the low back merger, I also present Pillai score data which more reliably quantify the degree of merger for individual participants by measuring vowel overlap (Nycz & Hall-Lew 2013). Though only certain vowels are analyzed here, the full list of F1 and F2 means and standard deviations for each vowel by region is included in Appendix C. The full ANOVA and Tukey test results for formant values and Euclidean distances are presented in Appendix E.

Gender could only be compared in Niagara, ON due to a shortage of male participants elsewhere but followed the same method as the regional comparison. Apparent time change was analyzed using Pearson correlations between various acoustic measures and year of birth, where a positive correlation indicates an increase as year of birth increases (=younger speakers).

3.3 Results

This section will outline the main results of the acoustic study: the demographic distribution (3.3.1), the low back merger and /æ/-tensing (3.3.2), the Low-Back-Merger Shift (3.3.3), allophones of /æ/ (3.3.4), /u:/-fronting and the Index of Phonetic Innovation (3.3.5), Canadian Raising (3.3.6), foreign (a) (3.3.7), the style comparison with spontaneous speech (3.3.8), and the qualitative attitude results (3.3.9). Each section will outline the regional results followed by any apparent time correlations and gender differences (limited to Niagara, ON).

3.3.1 Demographics

In total, 66 interviews were recorded for the acoustic study. The demographic distribution of participants is summarized in Table 16. It proved difficult to recruit a sufficient number of men from Niagara, NY and the GTA; gender differences will therefore only be analyzed for Niagara,

ON. Additionally, only women and men will be compared as there is only one participant who identified as another gender.

Region	**Niagara, ON**			**Niagara, NY**		**Greater Toronto Area**	
Gender	**F**	**M**	**O**	**F**	**M**	**F**	**M**
Participants	31	10	1	11	1	9	3
Total	42 (63.6%)			12 (18.2%)		12 (18.2%)	

Table 16: Summary of participants. F = female; M = male, O = other gender.

The participants' year of birth ranged from 1947 to 2001[7]. The mean was 1974 (approx. 48 years old) and the median was 1973 (approx. 49 years old). These participants therefore tended to be younger than in the questionnaire, though a wide range of ages is still covered. In Niagara, ON, this full range was found though participants were slightly younger than the overall group (range: 1947–2001, mean: 1976, median: 1975). As with the questionnaire, participants were younger in the GTA (range: 1953–1998, mean: 1979, median: 1983) and quite a bit older in Niagara, NY (range: 1949–1997, mean: 1962, median: 1957). Year of birth will be examined by individual region only, avoiding potential issues arising from these different distributions.

3.3.2 Low back merger and /æ/-tensing

One of the main expected differences between the Canadians and Americans is that the Canadians should merge the low back vowels LOT (/ɑ/) and THOUGHT (/ɔ/) and the Americans in this region should not. The low back merger contributes to the Low-Back-Merger Shift (to be discussed in 3.3.3) and the non-merger contributes to the Northern Cities Shift (NCS). The other major component of the NCS is the tensed TRAP vowel, where /æ/ is fronted and raised. If the border acts a barrier for these patterns, we expect the Canadians to have lower, unfronted /æ/ and merged /ɑ/-/ɔ/ while the Americans have tensed /æ/ and unmerged /ɑ/-/ɔ/. We must therefore look at these three vowels in each region (Figures 2-4).

[7] Year of birth was not provided for two participants from Niagara, ON.

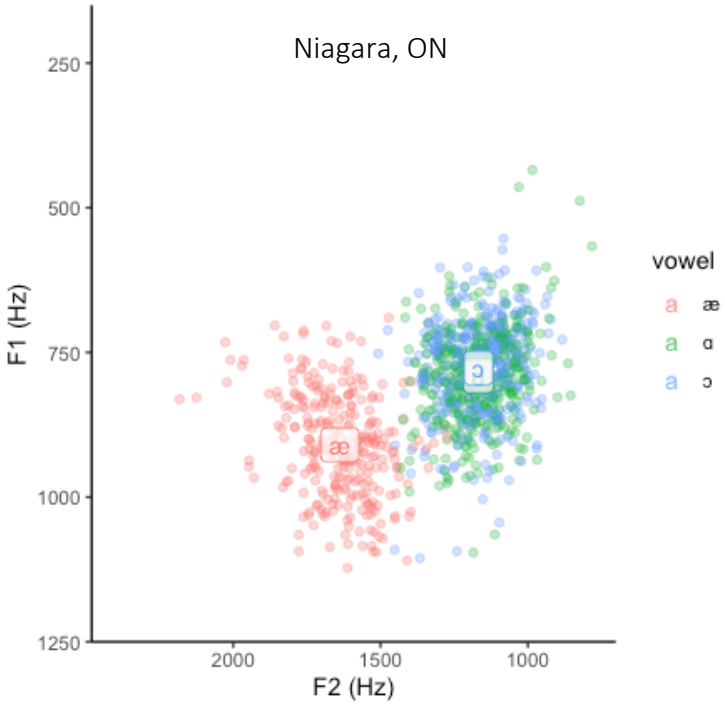

Figure 2: Individual tokens (dots) and means (labelled boxes) of /ɑ/, /ɔ/, and /æ/ in Niagara, ON.

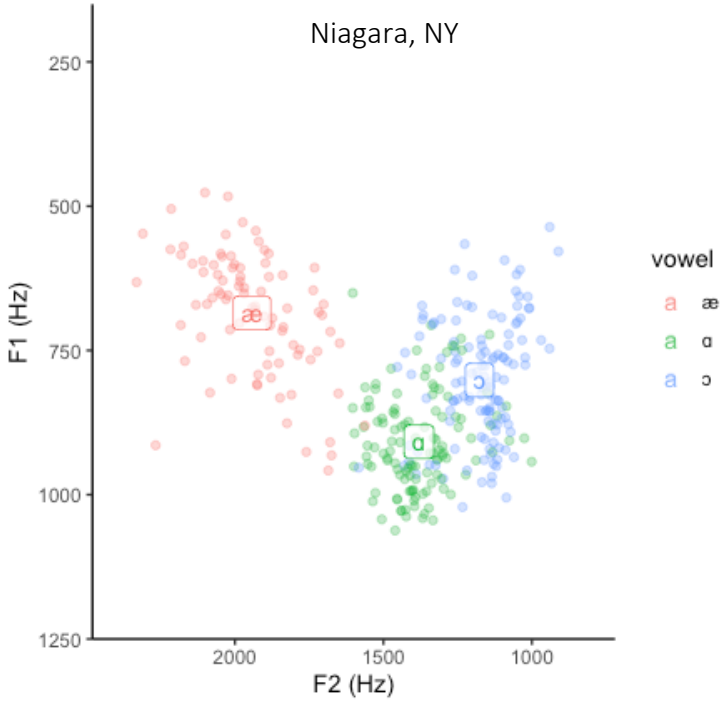

Figure 3: Individual tokens (dots) and means (labelled boxes) of /ɑ/, /ɔ/, and /æ/ in Niagara, NY.

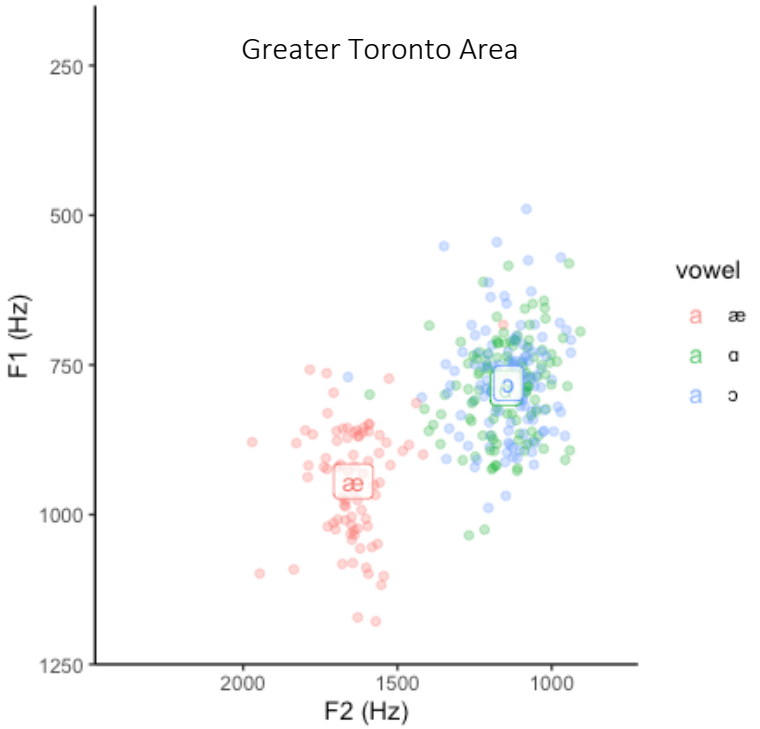

Figure 4: Individual tokens (dots) and means (labelled boxes) of /ɑ/, /ɔ/, and /æ/ in the Greater Toronto Area.

From these plots, it is clear that /æ/ is tensed in Niagara, NY but not in the Canadian regions. In Niagara, NY /æ/ is significantly higher and fronter than Niagara, ON (F1: $p<0.001$; F2: $p<0.001$) and the GTA (F1: $p<0.001$; F2: $p<0.001$). Niagara, ON and the GTA do not differ in fronting, but Niagara, ON /æ/ is significantly higher (F1: $p<0.05$). However, the difference in means is only 35Hz compared to the >200Hz differences between Canadians and Americans (see Appendix C).

It is also clear that both Canadian regions have merged /ɑ/-/ɔ/ and Niagara, NY does not. We can first look at this merger overall using ANOVAs. In Niagara, NY, /ɑ/ and /ɔ/ differ significantly in terms of F1 ($p<0.001$) and F2 ($p<0.001$). In contrast, /ɑ/ and /ɔ/ are not statistically different in F2 for Niagara, ON or the GTA. However, F1 is significantly higher for /ɑ/ than /ɔ/ in Niagara, ON ($p<0.05$). This may simply be due to a few outlier tokens of /ɑ/ which are particularly high and of /ɔ/ which are particularly low, as seen in Figure 2. Combined with the /æ/ difference above, this may show some diffusion of the American system; however, the differences between Niagara, ON and the GTA are still much smaller than any difference between Americans and Canadians.

Pillai scores allow us to further examine the presence of merger by measuring the overlap of /ɑ/ and /ɔ/ for individual speakers. These scores are then grouped by region (Figure 5). A higher Pillai score indicates more separation of the vowels (i.e., non-merger).

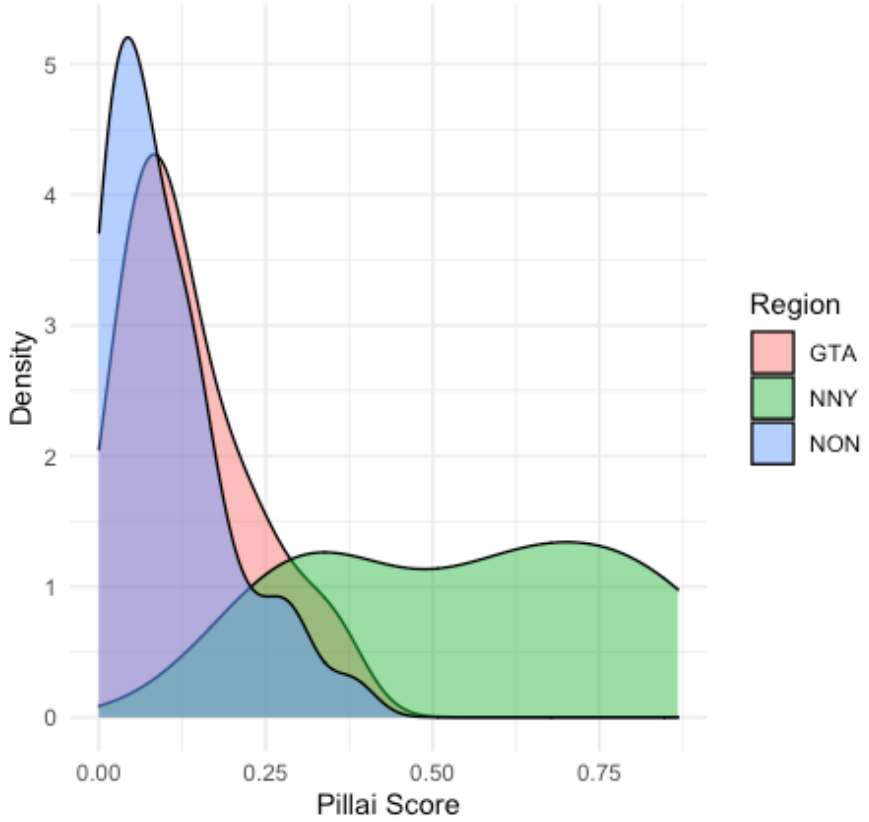

Figure 5: Comparison of Pillai scores by region for the /ɑ-/ɔ/ merger. 0 = fully merged, 1 = fully separated. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

From the Pillai scores, it is again immediately apparent that /ɑ/ and /ɔ/ are overwhelmingly merged in Niagara, ON and the GTA but not in Niagara, NY. Though there is a greater spread of scores in Niagara, NY, the proportion of speakers under 0.25 is low, especially compared to the vast majority of Canadians who are under 0.25. The scores immediately above 0.25 are of particular interest however: there are Canadians from both regions who fall into this area, along with many of the Americans. This suggests that some Canadians may not completely merge /ɑ/-/ɔ/ and some Americans may partially merge them. Whether this is actually salient would require further study.

Year of birth is correlated with Pillai scores in Niagara, NY: there is a decrease in Pillai scores for younger speakers that is fairly strong (r=-0.68, p<0.05). This could indicate merger in progress though, as seen in Figure 5, nearly no American speakers show complete merger at this

point. In Niagara, ON and the GTA, there is no significant correlation. The gender analysis in Niagara, ON also shows no significant effect of gender. If any of the Canadian speakers with higher Pillai scores do noticeably separate the vowels, this likely reflects speaker variation rather than an unexpected de-merger in progress.

3.3.3 Low-Back-Merger Shift

The Low-Back-Merger Shift (LBMS) involves retraction and/or lowering of the KIT (/ɪ/), DRESS (/ɛ/), and TRAP (/æ/) vowels. To see the relative position of these vowels, we can compare them to a point that does not move, namely the high front FLEECE (/i:/) vowel. I will follow Boberg (2019a, 2019b, 2021) in combining the three vowels using the LBMS index. This index is the mean of the mean Euclidean distances between /i:/ and each of /ɪ/, /ɛ/, and /æ/. I thus calculated the LBMS index for each region for comparison.

I first calculated the Euclidean distance between each speaker's mean /i:/ and each of the LBMS vowels. The mean across speakers was then calculated to give the regional mean for each vowel and for the three vowels together (LBMS index). The mean distance of individual vowels and the LBMS index for each region are presented in Table 17.

	Region		
Measure	**Niagara, ON**	**Niagara, NY**	**Greater Toronto Area**
/i:/ – /ɪ/	598 *(SD=142)*	599 *(SD=56)*	627 *(SD=126)*
/i:/ – /ɛ/	895 *(SD=143)*	872 *(SD=92)*	922 *(SD=143)*
/i:/ – /æ/	1138 *(SD=177)*	690 *(SD=163)*	1170 *(SD=132)*
LBMS Index	877 *(SD=270)*	720 *(SD=159)*	906 *(SD=260)*

Table 17: Mean Euclidean distances (Hz) for each LBMS vowel and overall (=LBMS index), rounded to the nearest whole number. SD = standard deviation (rounded to the nearest whole number).

It is immediately apparent that both Canadian regions have a higher LBMS index than Niagara, NY. This is confirmed statistically: Niagara, NY differs from both Niagara, ON (p<0.01) and the GTA (p<0.01). Niagara, ON and the GTA do not differ. This is as expected if the border separates the Low-Back-Merger Shift in Canada from the Northern Cities Shift in New York. Note that the difference in the LBMS index is driven by the /i:/ – /æ/ distance since /ɪ/ and /ɛ/ are also retracted and/or lowered in the NCS. As seen in Figure 6, the means of /ɪ/ and /ɛ/ are quite close between all three regions but the mean of /æ/ is low and retracted in Canada and tensed in Niagara, NY.

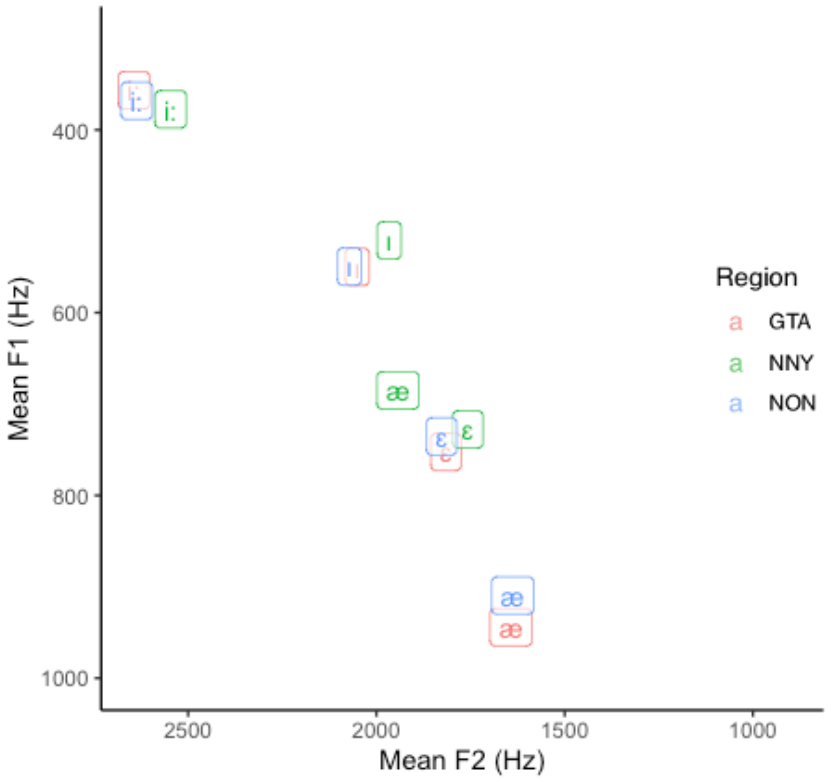

Figure 6: Mean formant values for LBMS vowels (/ɪ/, /ɛ/, /æ/) and the comparison point /i:/ in each region. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

The gender analysis in Niagara, ON showed women to have a higher LBMS index (mean: 919Hz) than men (mean: 756Hz) (p<0.05). This suggests women are further ahead in the shift, in line with previous findings (e.g., Clarke et al. 1995; Boberg 2010, 2019a, 2019b). However, there are no significant correlations between the LBMS index and year of birth in any of the regions. There are also no significant year of birth correlations for the Euclidean distances of individual vowels, though /ɪ/ in Niagara, ON (r=-0.3, p=0.055) is close. This correlation would be the opposite of what we expect if the vowels are continuing in the direction of retraction. Together with the mostly insignificant results, this may suggest the LBMS is slowing down. There is however greater variance in the Canadian regions (see SDs in Table 17) which may reflect the more recent progression of this shift. In any case, the Canadians clearly remain distinct from the Americans.

3.3.4 Allophones of /æ/

The raising and fronting of /æG/ and /æN/ can be measured by comparing their position to that of /æ/ (i.e., in other environments). I have already shown that /æ/ is low and retracted for the Canadians (3.3.3) and high and fronted for the Americans (3.3.2). We expect only /æG/ and /æN/

to be raised for the Canadians, meaning there will be a large distance between these allophones and /æ/. Further, we expect /æN/ to be the most tensed since this is the pattern found in Ontario (Boberg 2010). For Niagara, NY, all three should be raised and /æN/ may show additional raising and fronting. In Figure 7, these expectations are confirmed: in both Canadian regions (blue and red), /æN/ is highest followed by /æG/ while /æ/ remains low and retracted. In Niagara, NY, /æG/ is in line with tensed /æ/ but /æN/ is higher and fronter. Note that there is also a difference in quality with the NY /æ/ being diphthongized (Labov et al. 2006); this is not visible here as only nuclei were measured, but impressionistic analysis suggests that Canadians do not adopt this pronunciation in either region.

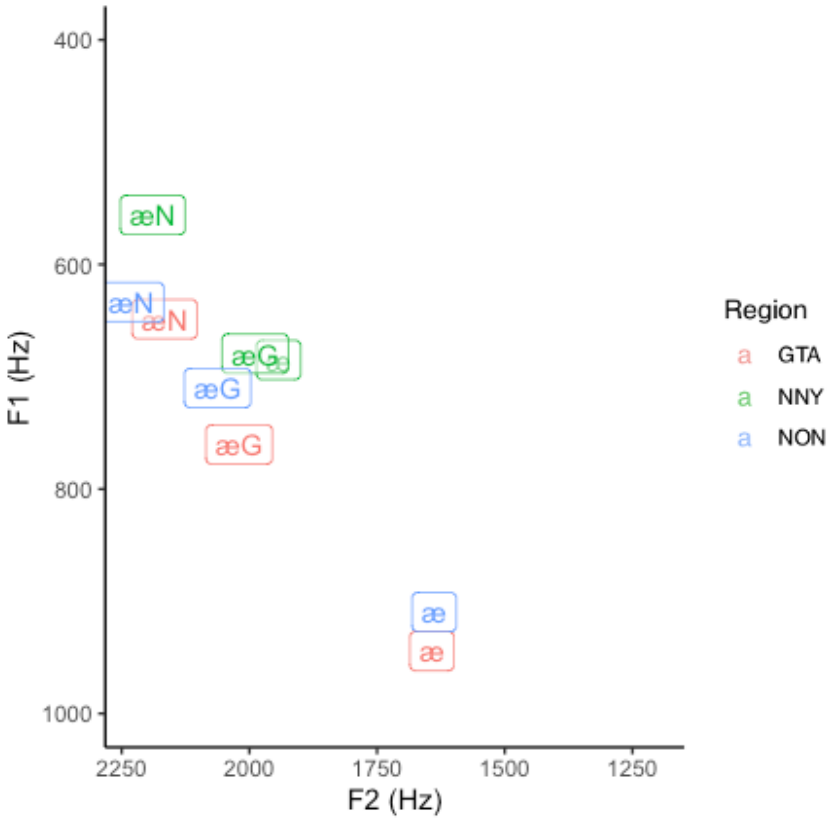

Figure 7: Mean formant values for /æ/, /æG/, and /æN/ in each region. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

These observations are confirmed statistically. In both Canadian regions, /æG/ is higher and fronter than /æ/ (both regions – F1: $p<0.001$, F2: $p<0.001$) while /æN/ is higher and fronter than both /æ/ (both regions – F1: $p<0.001$, F2: $p<0.001$) and /æG/ (both regions – F1: $p<0.001$, F2: $p<0.001$). In Niagara, NY, only /æN/ differs. It is significantly higher and fronter than both /æ/ (F1: $p<0.001$, F2: $p<0.001$) and /æG/ (F1: $p<0.001$, F2: $p<0.001$). Thus, Canadians only raise /æ/

before voiced velars and front nasals while Americans raise /æ/ everywhere, with both groups having the most raising and fronting before front nasals.

The regions can also be compared in terms of the Euclidean distances between each of the allophones:

	Region		
Measure	**Niagara, ON**	**Niagara, NY**	**Greater Toronto Area**
/æ/ – /æG/	473 *(SD=131)*	120 *(SD=66)*	425 *(SD=123)*
/æ/ – /æN/	658 *(SD=202)*	293 *(SD=119)*	623 *(SD=190)*
/æG/ – /æN/	205 *(SD=149)*	256 *(SD=126)*	225 *(SD=116)*

Table 18: Mean Euclidean distances (Hz) between each allophone of /æ/, rounded to the nearest whole number.

The difference in Euclidean distances is never significant between Niagara, ON and the GTA, suggesting that these regions have an equal degree of raising and fronting. Niagara, NY does have significantly lower Euclidean distances for both /æ/–/æG/ (NON: $p<0.001$, GTA: $p<0.001$) and /æ/–/æN/ (NON: $p<0.001$, GTA: $p<0.001$). However, /æG/–/æN/ does not differ between the Canadians and Americans, meaning /æN/ is equally more raised and fronted than /æG/ in all three regions. Notably, the variance for /æG/–/æN/ is quite large in all regions compared to the means; some speakers may therefore show no difference while others show a large difference.

The gender comparison in Niagara, ON shows women to have greater /æ/–/æG/ ($p<0.01$) and /æ/–/æN/ ($p<0.01$) distances than men, with no difference in /æG/–/æN/. This mainly reflects greater /æ/ retraction (F1: $p<0.001$, F2: $p<0.001$) for women, though there is also greater fronting of /æN/ (F2: $p<0.01$) for women. The only apparent time correlation is in Niagara, NY: younger speakers have a much greater /æ/–/æN/ distance ($r=0.65$, $p<0.05$). This particular difference may then be increasing while the Canadian pattern remains stable, though this may be more influenced by increasing retraction of /æ/ as we will see in the next section.

3.3.5 Index of Phonetic Innovation

The Index of Phonetic Innovation looks at the degree of /u:/-fronting and /æ/-retraction to give an overall value of how innovative speakers are (Boberg 2010, 2021). It is calculated by subtracting the mean F2 of /u:/ (GOOSE) from the mean F2 of /æ/ (TRAP). If /u:/ is very fronted and /æ/ very retracted the IPI will be negative (innovative speaker), whereas less movement of the vowels means the IPI will be positive (conservative speaker). As has already been shown, the Americans tense /æ/ while the Canadians retract /æ/, so the Canadians should have more innovative

scores. We will now also see how fronted /u:/ is in each region. The IPI was calculated for each speaker and speakers were compared by region.

As expected, Niagara, NY (mean: 494Hz, SD: 205) has significantly higher (i.e., less innovative) IPI values than both Niagara, ON (mean: -158Hz, SD: 197; $p<0.001$) and the GTA (mean: -228Hz, SD: 192; $p<0.001$). The Canadian regions did not significantly differ from each other. We additionally see that Niagara, NY has a very positive mean IPI indicating more conservative pronunciation while the Canadians have very negative mean IPIs indicating more innovation. This also means that /æ/ continues to be further forward than /u:/ for the Americans while it is retracted behind /u:/ for the Canadians. Figure 8 shows the mean /æ/ and /u:/ values in each region in relation to each other and to the high front (/i:/) and back (/u:l/) corners of the vowel space.

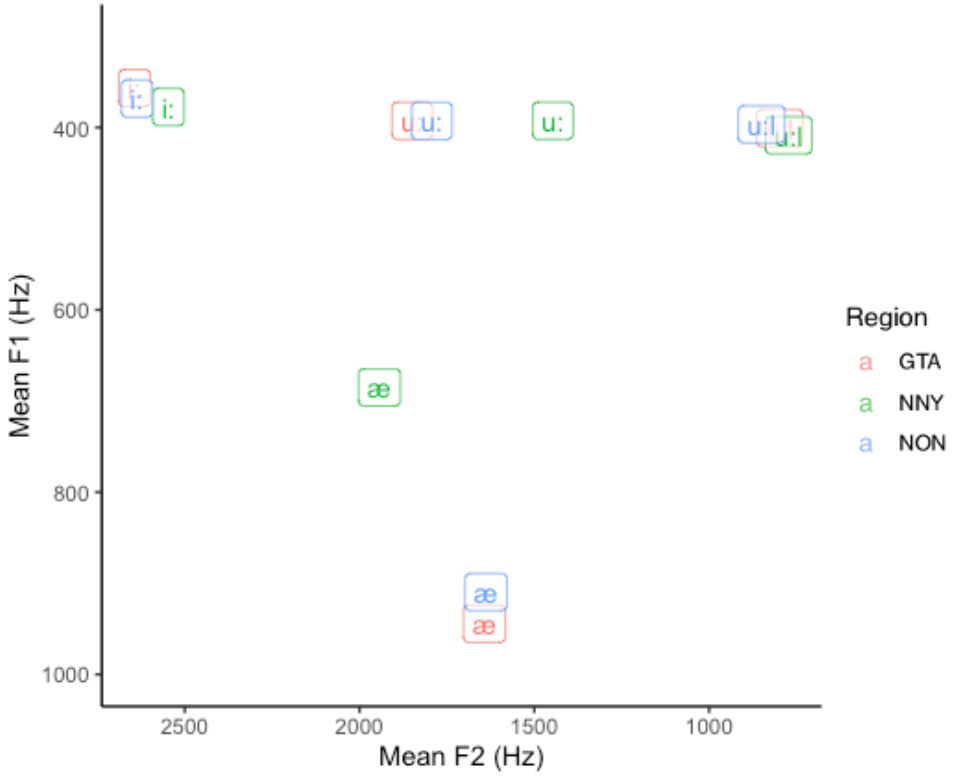

Figure 8: Mean formant values for /æ/ and /u:/ in each region with /i:/ and /u:/ as comparison points. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

In addition to /æ/ being more retracted for Canadians, we now see that /u:/ is more fronted than Niagara, NY in both Niagara, ON (F2: $p<0.001$) and the GTA (F2: $p<0.001$). There is no difference between Niagara, ON and the GTA, meaning Canadians have an equal degree of both /u:/-fronting and /æ/-retraction (see 3.3.3). This is reflected in their similar IPI values.

The only significant correlation with year of birth is in Niagara, NY: younger speakers have much lower IPI values (r=-0.61, p<0.05). Looking at the two values individually, this seems to be influenced more by /æ/: while there is no correlation for the F2 of /u:/, the F2 of /æ/ decreases with year of birth (r=-0.63, p<0.05). These apparent time findings suggest increasing retraction of /æ/ in Niagara, NY, though we would again need real time data to differentiate this from an age-grading effect. This pattern is however in line with a reversal of the NCS in younger speakers as has been previously reported (e.g., Driscoll & Lape 2015 in Syracuse, NY).

3.3.6 Canadian Raising

For Canadian Raising, we need to compare the tokens where the diphthong is open or followed by a voiced obstruent (/ai/, /au/) to the tokens where the diphthong is followed by a voiceless obstruent (/aiT/, /auT/; 'T' represents any voiceless obstruent). If there is Canadian Raising, the nucleus of the diphthong will be higher for /aiT/ and /auT/ than for their respective counterparts. Looking at the mean /ai/, /au/, /aiT/, and /auT/ formant values in each region, we see that in fact all three regions have some degree of both types of raising (Figure 9):

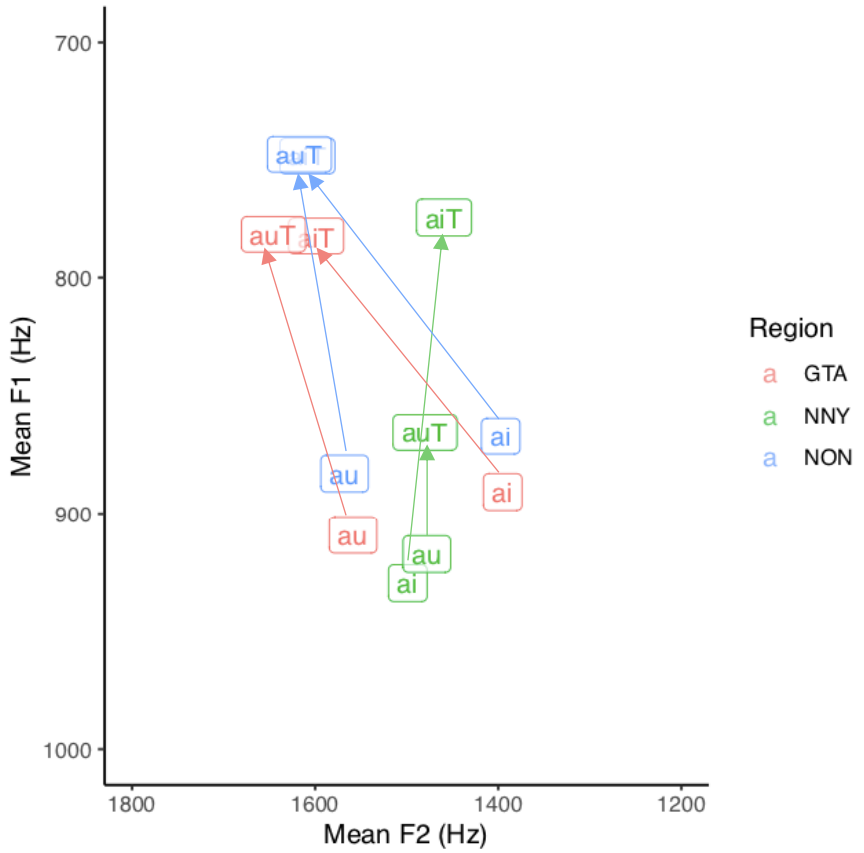

Figure 9: Mean formant values for Canadian Raising variables in each region. /au/ = MOUTH, /ai/ = PRICE; 'T' represents any voiceless obstruent following the diphthong (=expected raised tokens). GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

Raising is confirmed statistically in each case. In Niagara, ON both the /aiT/ and /auT/ tokens are significantly higher (F1; /au/-/auT/: $p<0.001$, /ai/-/aiT/: $p<0.001$) and more fronted (F2; /au/-/auT/: $p<0.01$, /ai/-/aiT/: $p<0.001$) than their counterparts. The GTA shows the exact same pattern of raising (F1; /au/-/auT/: $p<0.001$, /ai/-/aiT/: $p<0.001$) and fronting (F2; /au/-/auT/: $p<0.01$, /ai/-/aiT/: $p<0.001$). Additionally, /au/ is further forward than /ai/ in both regions (NON: $p<0.001$, GTA: $p<0.001$), but there is no difference between /auT/ and /aiT/. The nucleus of both diphthongs is therefore fronted to approximately the same position despite unraised /ai/ being further back.

In Niagara, NY, raising is limited to height: both /aiT/ and /auT/ are higher than their counterparts (F1; /au/-/auT/: $p<0.05$, /ai/-/aiT/: $p<0.001$) but are not further forward. While /ai/-raising is not particularly surprising, /au/-raising is not as common in the U.S. However, while the difference (51Hz) is significant it does not reach the minimum difference for raising (60Hz) set out in the Atlas of North American English (Labov et al. 2006).

We can again combine the F1 and F2 dimension of individual speaker means as Euclidean distances to compare the degree of raising. The distances between /ai/-/aiT/ and /au/-/auT/ were calculated for each speaker then compared across regions. For /ai/-/aiT/, Niagara, ON (mean: 253Hz, SD: 93Hz) has significantly greater distances than Niagara, NY (mean: 174Hz, SD: 53Hz; $p<0.05$); no other differences were significant (GTA – mean: 243Hz, SD: 85Hz). For /au/-/auT/, Niagara, NY (mean: 77Hz, SD: 52Hz) had significantly lower distances than both Niagara, ON (mean: 172Hz, SD: 80Hz; $p<0.01$) and the GTA (mean: 182Hz, SD: 104Hz; $p<0.01$). Raising of /aiT/ in Niagara, NY is therefore equivalent to Toronto while /auT/ is raised to a much lower degree than both Canadian groups. However, the Euclidean distance mean (77Hz) does reach Labov et al.'s (2006) threshold for Canadian Raising.

The Euclidean distances show no correlations with year of birth. However, the individual F1 and F2 of /aiT/ and /auT/ show some patterns. In Niagara, ON, /auT/ shows increased fronting for younger speakers (F2: $r=0.48$, $p<0.01$) but there is no change in height and /aiT/ shows no change. In the GTA, there is no significant correlation for either /auT/ or /aiT/, though both are close to showing a decrease in height (increase in F1; /auT/: $r=0.53$, $p=0.077$; /aiT/: $r=0.5$, $p=0.097$); a greater number of participants would show whether this is spurious or a possible change in the degree of raising. Niagara, NY also shows no significant correlations; /au/-raising

therefore does not seem to be an emerging pattern in NY and may instead be linked to speaker variation, though more participants would also clarify this pattern.

For Canadian Raising, we are interested not only in whether there is diffusion but in what ways the pattern might be changing. As mentioned, I impressionistically noticed raising in *houses* and *mouths* despite the following voiced /z/. I therefore analyzed these words separately. In Figure 10, the /au/ in *houses* is in fact noticeably raised in both Canadian regions.

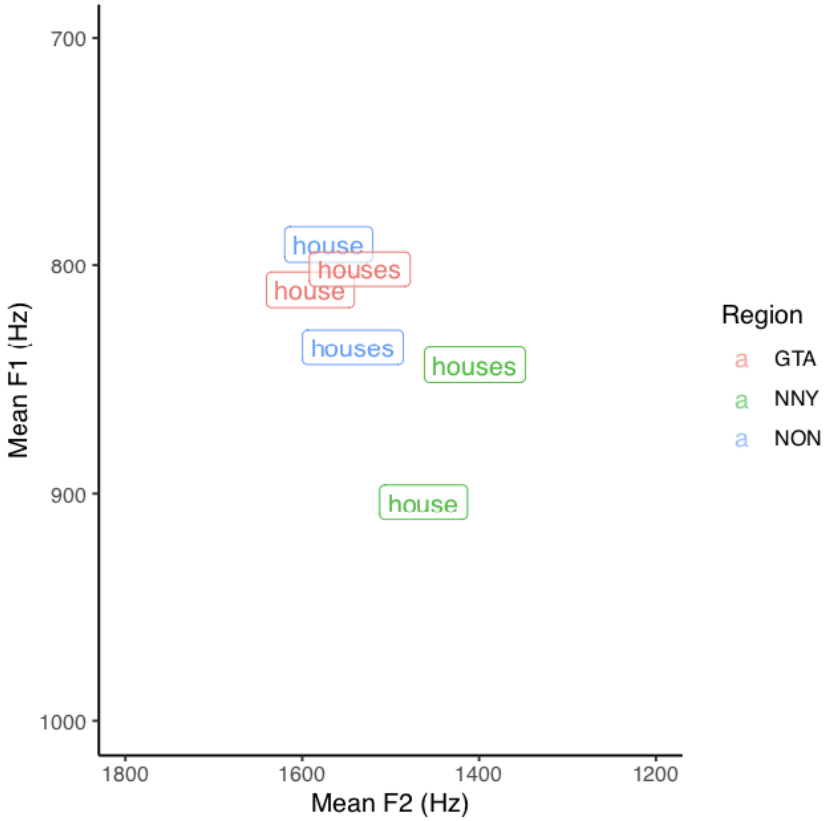

Figure 10: Mean formant values for *house* (/auT/) and *houses* (/au/) in each region. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

The GTA very clearly shows this pattern with *house* and *houses* overlapping; the height difference between them is non-significant (F1: p=0.8). They are also quite close in Niagara, ON though the difference is still marginally significant (F1: p=0.0465); however, *houses* is still higher (mean F1: 835Hz) than their unraised /au/ (mean F1: 892Hz, when *houses* is excluded). There is also an odd pattern in Figure 10 for Niagara, NY: *houses* is actually much higher than *house*. There is no clear explanation for this given that it is opposite to the overall /au/-raising pattern noted earlier. It may however contribute to the small raising difference we saw in Figure 9: the mean

/au/-/auT/ distance in F1 increases from 52Hz to 72Hz when *house* and *houses* are excluded, which then meets Labov et al.'s (2006) raising threshold like the Euclidean distance does.

The results for *mouths*, shown in Figure 11, do not show the same degree of /au/-raising as *houses*. Though *mouths* appears to be slightly higher in Niagara, ON, it is nowhere near the height of *mouth* nor *houses*.

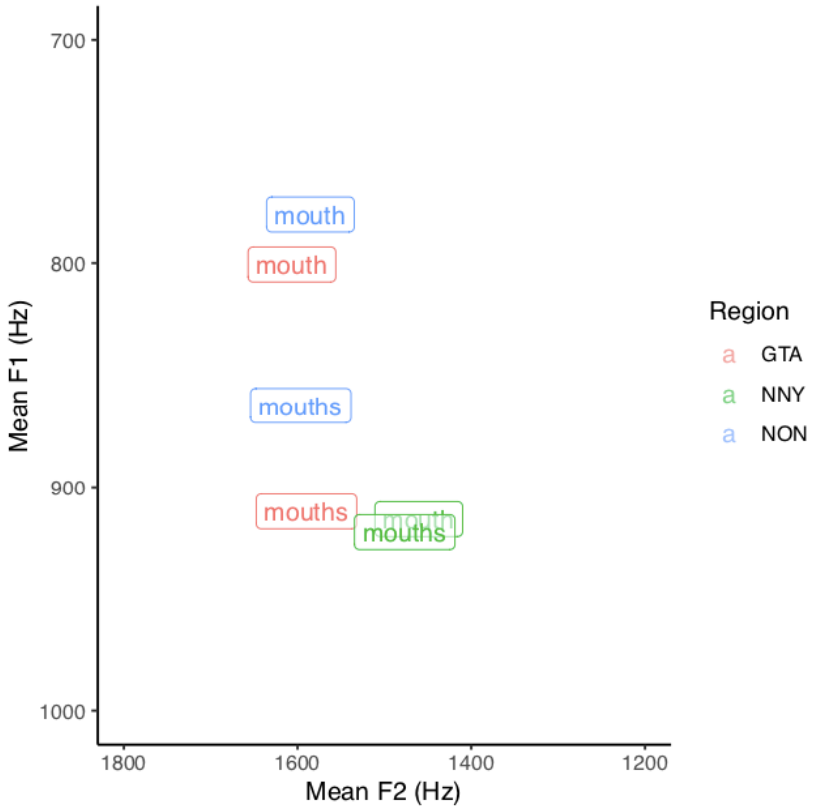

Figure 11: Mean formant values for *mouth* (/auT/) and *mouths* (/au/) in each region. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

However, this actually seems to be because it is only raised for a subset of speakers in Niagara, ON: if we take the mean of *houses* (F1: 836Hz) as a minimum for raising, 16 of the 37 *mouths* tokens in Niagara, ON are actually raised (F1 range: 640Hz – 834Hz). Raising in *mouths* may therefore simply be more variable than in *houses*. Interestingly, there are no apparent time correlations for *houses*-raising or *mouths*-raising. There is also no gender difference. These instances of /au/-raising may therefore reflect speaker variation rather than change in progress.

3.3.7 Foreign (a)

The acoustic foreign (a) data show that, as with the questionnaire data, Niagara, ON is different from the GTA, but only in terms of height (F1: $p<0.01$). The acoustic data also reveal that Niagara, ON is not actually similar to Niagara, NY either. In fact, Niagara, ON and Niagara,

NY are the most different from each other if we compare the overall mean F1 and F2 values across foreign (a) variables:

	F1	Difference (Hz)	F2	Difference (Hz)
NON-GTA	**	-33	-	-42
NON-NNY	***	-58	***	-119
GTA-NNY	-	-25	*	-77

Table 19: Overall differences between each region in F1 and F2 for all foreign (a) variables. The difference in normalized Hertz between means of the first and second region in each pair is indicated. * = p<0.05, ** = p<0.01, *** = p<0.001. GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

The mean F1 and F2 values in each region further reveal that the pronunciation of (a) in Niagara, ON tends to be more back (lower F2) and higher (lower F1). The mean (a) pronunciations for each region are shown together in Figure 12.

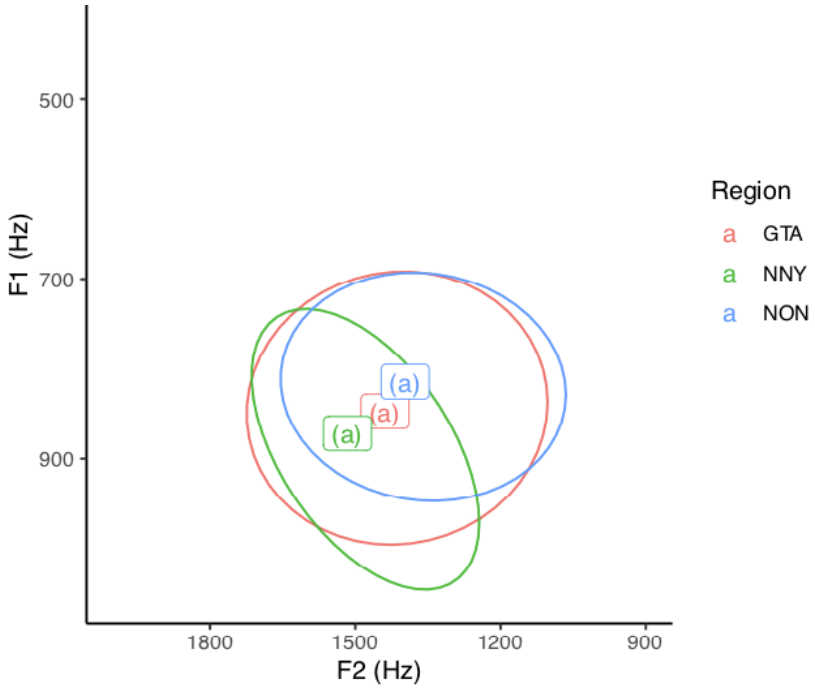

Figure 12: Mean foreign (a) formant values (boxes) in each region plotted with a 68% confidence interval (ellipses). GTA = Greater Toronto Area; NNY = Niagara, NY; NON = Niagara, ON.

While the means show us the overall tendency for speakers to use a more back or front pronunciation, they also obscure the divide in which vowel is actually being used. The means seem to suggest that speakers in each region use a low central vowel, when in fact we expect some pronunciations to have a more front (a) vowel and others to have a more back (a) vowel. To see the actual vowel quality being used, we can compare the means and individual points to the three

low vowels: TRAP – /æ/, LOT – /ɑ/, and THOUGHT – /ɔ/. Recall that the Canadian regions have merged /ɑ/-/ɔ/ while the Americans have unmerged, fronted /ɑ/. In Figures 13-15, we see how the (a) means and individual datapoints align with /æ/, /ɑ/, and /ɔ/ in each region.

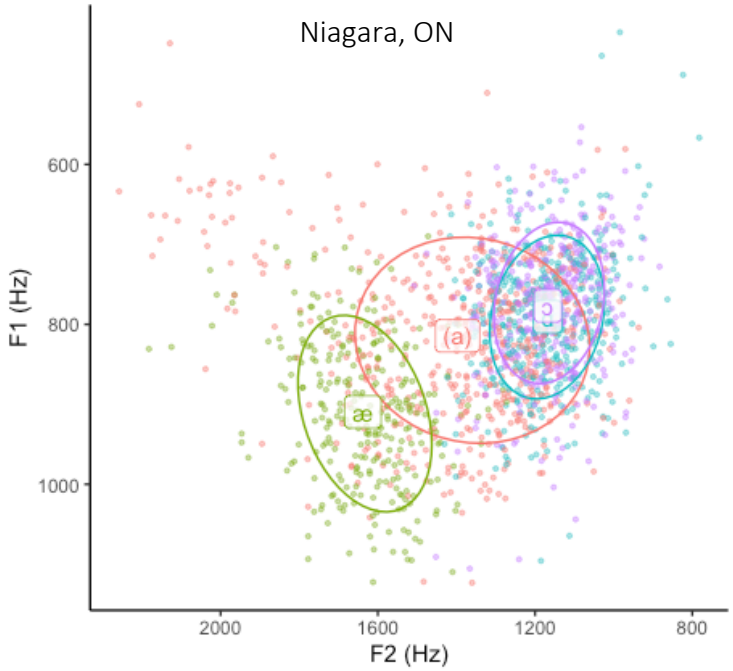

Figure 13: Comparison of mean and individual formant values of foreign (a), TRAP – /æ/, LOT – /ɑ/, and THOUGHT – /ɔ/ in Niagara, ON with a 68% confidence interval (ellipses).

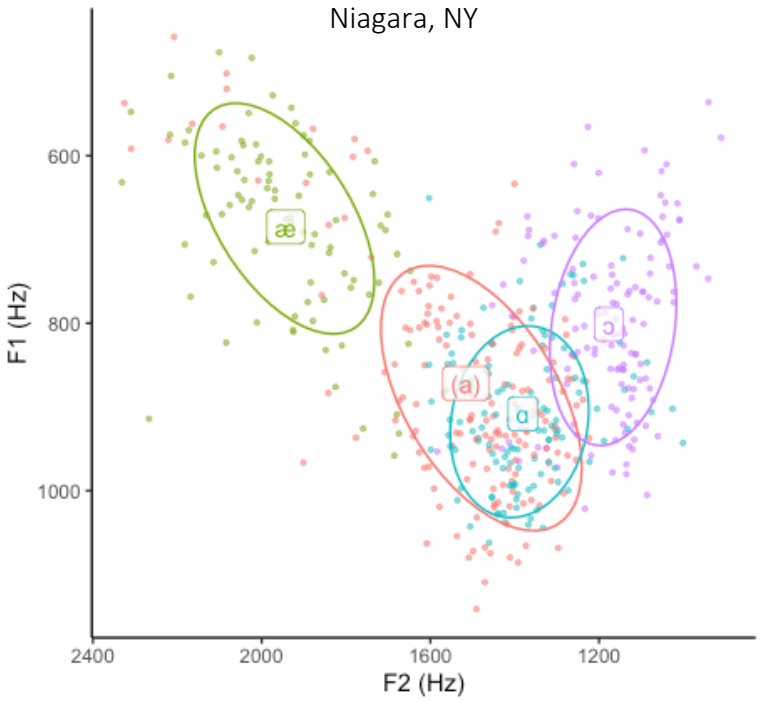

Figure 14: Comparison of mean and individual formant values of foreign (a), TRAP – /æ/, LOT – /ɑ/, and THOUGHT – /ɔ/ in Niagara, NY with a 68% confidence interval (ellipses).

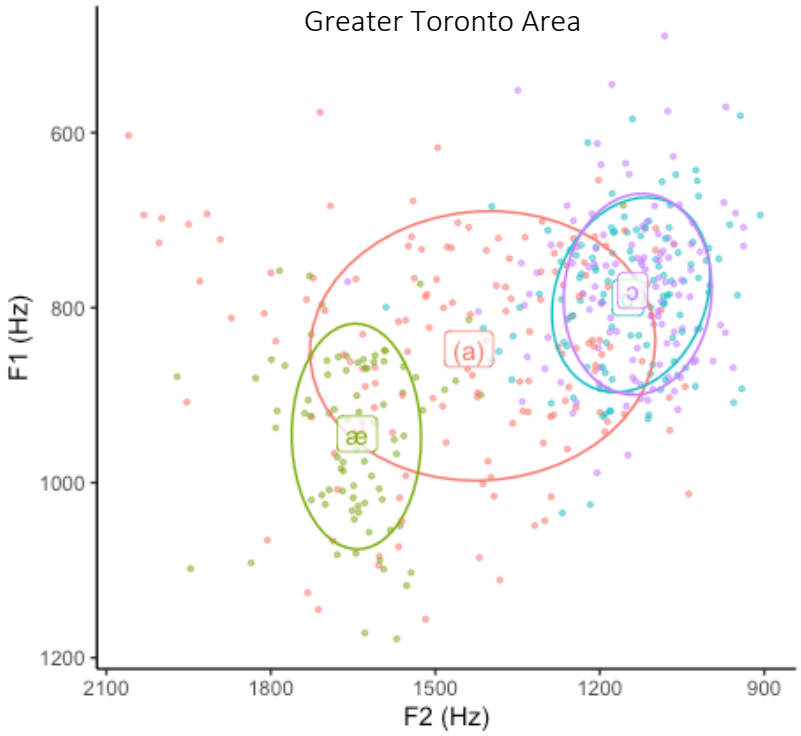

Figure 15: Comparison of mean and individual formant values of foreign (a), TRAP – /æ/, LOT – /ɑ/, and THOUGHT – /ɔ/ in the Greater Toronto Area with a 68% confidence interval (ellipses).

These figures make it clear that the pronunciation of (a) in both Canadian regions does not cluster around the mean. Instead, there is quite a bit of variance in (a)'s pronunciation, especially in the GTA. The individual tokens are spread throughout the range between TRAP and LOT-THOUGHT with many overlapping the /æ/ and /ɑ/-/ɔ/ areas in the vowel space. The difference between Niagara, ON and the GTA is that a greater proportion of the Niagara (a) tokens are clustered in the /ɑ/-/ɔ/ region towards the back and consequently a greater proportion of the (a) ellipse (=one standard deviation) overlaps with the /ɑ/ and /ɔ/ ellipses.

On the other hand, the (a) tokens in Niagara, NY generally cluster around the mean and overwhelmingly overlap with the /ɑ/ region. A few of the tokens are instead more closely aligned with the tensed /æ/, which we will see are word-specific; however, nearly no (a) tokens appear in the /ɔ/ region. The back (a) pronunciation in Niagara, NY is therefore distinct from the Canadian back (a) since it is primarily linked to the fronted LOT vowel which the Canadians do not have.

In all cases however, there are at least some datapoints corresponding to both the TRAP vowel and the LOT or merged LOT-THOUGHT vowel. The main difference between the two vowels is in the backness dimension with TRAP being further forward (higher F2), and an additional height difference for Niagara, NY (lower F1). If we look at the mean formant values in each region by

word, it becomes clear that certain words are more likely to be front or back. The means of individual foreign (a) variables were plotted in each region with the range (95% confidence interval) of /æ/, /ɑ/, and /ɔ/. A table with the full list of foreign (a) means by word is provided in Appendix D. In Niagara, ON (Figure 16), it is clear that most (a) variables tend to be back:

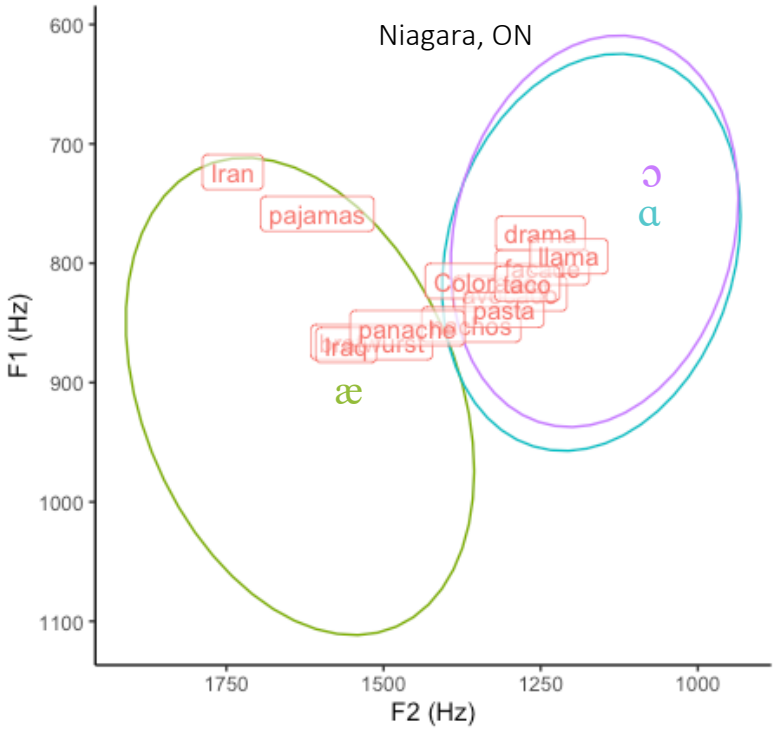

Figure 16: Mean formant values of (a) for individual words in Niagara, ON, with 95% confidence intervals of /æ/, /ɑ/, and /ɔ/ (ellipses) for comparison.

Many words fall in the /ɑ/-/ɔ/ range: *drama, llama, façade, taco, avocado, pasta, Colorado*. Others clearly overlap with /æ/ instead: *Iran, Iraq, pajamas, bratwurst, panache*. Note that *Iran* and *pajamas* are particularly far forward and high due to the following nasal, as seen with the /æN/ variables in section 3.3.4. Since the 95% confidence intervals of /ɑ/-/ɔ/ and /æ/ overlap slightly, there is no clear space for intermediate (central) pronunciations of (a) since all of the words fall into one category. Though *nachos* is quite close to the centre, it still falls within the /ɑ/-/ɔ/ confidence interval.

In Niagara, NY (Figure 17), the individual variables are more consistently within the range of /ɑ/. Additionally, even though the range of /ɑ/ and /ɔ/ overlap, the (a) words tend to cluster towards the front of the /ɑ/ range away from /ɔ/.

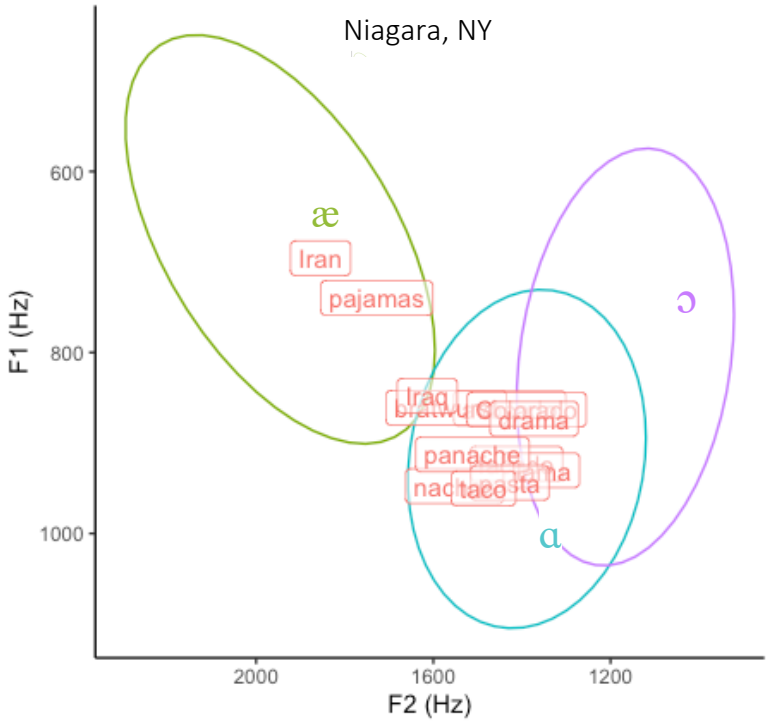

Figure 17: Mean formant values of (a) for individual words in Niagara, NY, with 95% confidence intervals of /æ/, /ɑ/, and /ɔ/ (ellipses) for comparison.

Like Niagara, ON, *Iran* and *pajamas* are quite far forward and raised to tensed /æ/; the fact that their means are below the /æN/ mean (F1: 556Hz) and the variance is large (Iran – F1: SD=155Hz, F2: SD=317Hz; pajamas – F1: SD=179Hz, F2: SD=311Hz) suggests at least some speakers instead use /ɑ/. *Iraq* also falls along the boundary of /æ/ and /ɑ/, suggesting at least some speakers use the /æ/ pronunciation. The split in Niagara, NY is therefore between /ɑ/ in the majority of words and /æ/ in at least a couple words, but never /ɔ/.

The pattern in the GTA is more similar to that in Niagara, ON, with a greater split between /æ/ and /ɑ/-/ɔ/ (Figure 18). Some words are clearly more front (*Iran, Iraq, bratwurst, panache, pasta*) or back (*taco, drama, avocado, façade, llama*). However, the range of /æ/ and /ɑ/-/ɔ/ do not overlap here, leaving more variables between the two vowels: *nachos*, *Colorado*, and *pajamas* are not clearly in either.

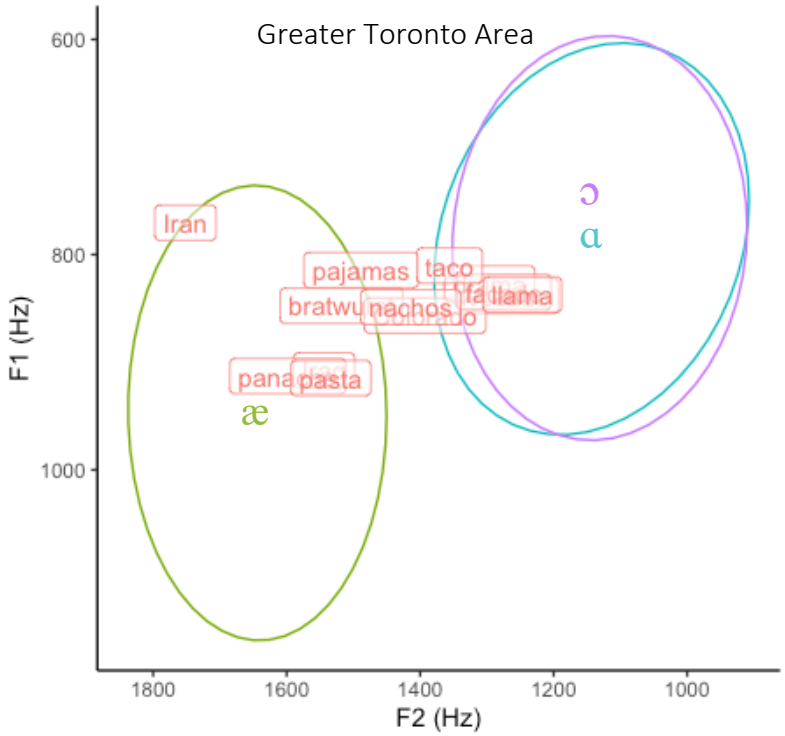

Figure 18: Mean formant values of (a) for individual words in the Greater Toronto Area, with 95% confidence intervals of /æ/, /ɑ/, and /ɔ/ (ellipses) for comparison.

To look more closely at these splits and the interregional differences, we can categorize individual tokens as /ɑ/, /æ/, or intermediate in relation to the speaker's mean /ɑ/ and /æ/ pronunciations. To do this, I follow one of Boberg's (2020) methods of foreign (a) categorization: tokens that fall within two standard deviations of the speaker's mean /ɑ/ or /æ/ are categorized as /ɑ/ or /æ/, respectively, and tokens that fall between the two are categorized as intermediate. If the ranges overlapped, I also categorized values within the overlap as intermediate. The percentages of each pronunciation for each word are summarized in Table 20 by region; the most frequent pronunciation in each region is underlined. Note that frequencies are rounded to the nearest whole number so some totals may not be exactly 100%.

	Niagara, ON			**Niagara, NY**			**GTA**		
Word	**/ɑ/**	**/æ/**	**inter.**	**/ɑ/**	**/æ/**	**inter.**	**/ɑ/**	**/æ/**	**inter.**
avocado	<u>63</u>	5	32	<u>92</u>	0	8	<u>58</u>	17	25
bratwurst	10	<u>68</u>	23	<u>62</u>	31	8	8	<u>67</u>	25
Colorado	41	15	<u>44</u>	<u>92</u>	0	8	<u>45</u>	27	27
drama	<u>80</u>	7	12	<u>92</u>	0	8	<u>58</u>	25	17
façade	<u>85</u>	5	10	<u>92</u>	0	8	<u>67</u>	17	17

Iran	20	66	15	23	69	8	17	75	8
Iraq	7	83	10	46	46	8	17	75	8
llama	83	5	12	92	0	8	64	9	27
nachos	27	20	54	50	8	42	17	25	58
pajamas	15	54	32	38	62	0	8	42	50
panache	32	46	22	62	8	31	0	75	25
pasta	70	18	12	85	0	15	9	73	18
taco	83	7	10	67	0	33	50	33	17
Mean	47	31	22	69	17	14	32	43	25

Table 20: Frequencies (%) of /ɑ/, /æ/, and intermediate (inter.) pronunciations of each foreign (a) variable by region, rounded to the nearest whole number.

This analysis confirms that intermediate pronunciations are relatively infrequent, though the means show these pronunciations make up 17-25% of tokens. Use of intermediate forms largely depends on the variable: *Colorado* and *nachos* actually favour intermediate pronunciations in Niagara, ON and *nachos* and *pajamas* favour intermediate pronunciations in the GTA. *Nachos* also has a high frequency of intermediate forms in Niagara, NY, suggesting this variable is more likely to be intermediate in general. This analysis of individual speakers also confirms that Niagara, ON has higher frequencies of /ɑ/ pronunciations than the GTA overall and specifically for several words (e.g., *drama, façade, llama, panache, pasta, taco*).

We again see the biggest difference in the pronunciation of *pasta*. The frequencies confirm that Niagara, ON speakers are overwhelmingly using their back /ɑ/-/ɔ/ vowel (70%) and not an intermediate form. In the GTA, we instead see speakers overwhelmingly using front /æ/ (73%). The pronunciation of *pasta* in each region is also plotted in Figure 19. Niagara, ON and the GTA differ significantly in both F1 ($p<0.05$) and F2 ($p<0.001$). Niagara, ON and Niagara, NY also differ in F1 ($p<0.001$) and are marginally non-significantly different in F2 ($p=0.055$). In both cases, Niagara, ON has a higher and relatively more back pronunciation. On the other hand, the GTA does not differ significantly from Niagara, NY in either dimension. This is likely due to the back (a) tokens (9%) and intermediate tokens (18%) in the GTA pulling the mean back; the two regions evidently do not use the same vowel in most cases, as shown in Table 20. The GTA mean and range is quite far forward whereas Niagara, ON tends towards the back. The 68% confidence interval is clearly larger in both Canadian regions due to some degree of both TRAP and LOT-THOUGHT use, whereas the Americans consistently use their LOT vowel.

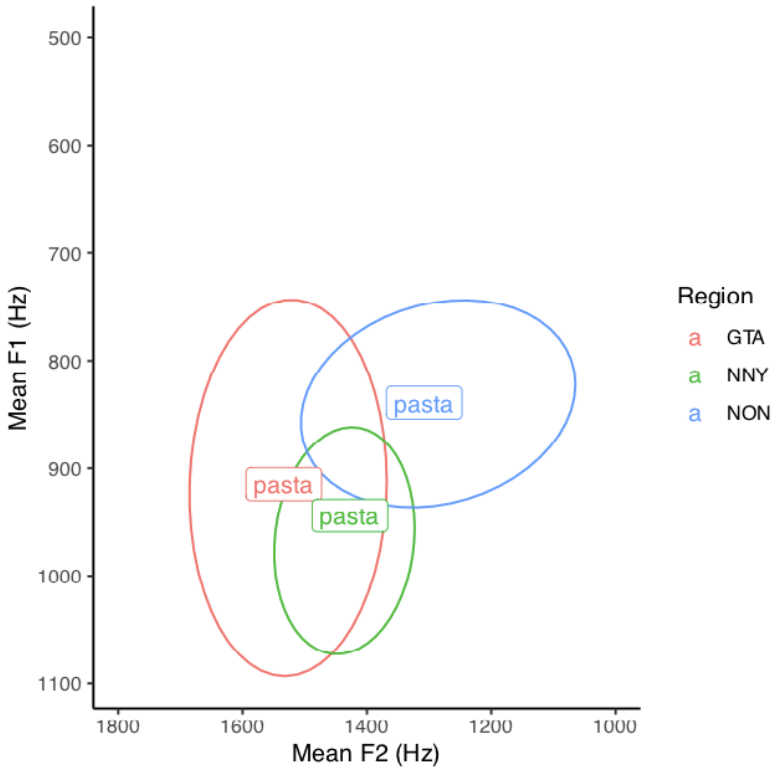

Figure 19: Mean formant values of (a) in *pasta* (boxes) with 68% confidence interval (ellipses) by region. GTA = Greater Toronto Area; NON = Niagara, ON; NNY = Niagara, NY.

The apparent time analysis of formant values shows that overall (a) in Niagara, ON is slightly further back for younger speakers (F2: r=-0.32, p<0.05); there is no significant correlation for height. Additionally, there is more backing of (a) in *pasta* specifically for younger speakers (NON: r=-0.41, p<0.05). Neither dimension in the GTA or Niagara, NY was correlated with year of birth for overall (a) or *pasta*, suggesting (a) is largely stable; however, a larger sample of these speakers may reveal a pattern. The gender analysis of overall (a) in Niagara, ON showed a significant difference in F2 (p<0.05) with women having a more back pronunciation (mean: 1380Hz) than men (mean: 1444Hz), suggesting women are ahead in the backing of (a).

3.3.8 Style comparison

The results reported above focus only on the word list data; the spontaneous speech will not be analyzed in full here. However, the two styles can be preliminarily compared to see if the results might carry over to spontaneous speech. In this section, I will compare the word list and spontaneous data for six of the Niagara, ON speakers (three born before 1980: A, D, F; three born 1980 or later: B, C, E), focusing on the vowels investigated above. The differences between speech styles are summarized for each speaker and vowel in Table 21.

speaker	A		B		C		D		E		F	
vowel	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
/ɑ/	-	-	* <	-	-	* >	-	-	* <	-	-	-
/ɔ/	-	-	* <	-	-	* >	-	-	-	-	-	-
/ɪ/	-	* <	-	-	-	-	-	* <	* <	* <	-	* <
/ɛ/	-	* <	-	-	-	-	-	* <	-	* <	-	-
/æ/	-	-	* <	-	* <	-	-	-	* <	-	-	-
/æG/			-	-	-	-						
/æN/	-	* <	-	-	-	-	* <	-	-	* <	* >	-
/u:/	-	-	-	-	-	-	-	-	-	-	-	-
/ai/	-	-	* <	* >	-	* >	-	-	-	-	-	-
/aiT/	-	-	-	-	* <	-	* <	-	* <	-	-	-
/au/	-	-	-	-	* <	-	-	* >	-	-	* <	-
/auT/	-	* <	-	-	-	-	-	-	-	-	* <	-
(a)	-	-					-	-	* <	-	-	-

Table 21: Differences between word list and spontaneous speech style for each vowel of interest for six Niagara, ON speakers. Significant differences are marked with * and the direction is indicated with < (spontaneous value lower than word list) and > (spontaneous value higher). Greyed out boxes indicate an insufficient number of tokens.

In general, there are more similarities than differences between the two speech styles. However, there are several differences to address. First, the low back merger vowels (/ɑ/ and /ɔ/) vary in height for two speakers and backness for one speaker. For Speakers B and C, the difference is in the same direction for both vowels suggesting they are still merged in spontaneous speech but higher and fronter, respectively. For Speaker E, only /ɑ/ is higher; however, /ɑ/-/ɔ/ is still merged as the two vowels for this speaker do not significantly differ in the spontaneous data. The word list merger results from 3.3.2 therefore seem to be reliable.

For the Low-Back-Merger Shift vowels (/ɪ/, /ɛ/, /æ/) there are also a few differences. For three of the six speakers /ɪ/ and/or /ɛ/ have lower F2 (more back), meaning the vowels are actually more retracted in spontaneous speech for some speakers. However /æ/ for three speakers and /ɪ/ for one speaker have lower F1 (higher), suggesting the degree of lowering could be exaggerated in the word list style data in 3.3.3.

On the other hand, the raising/fronting of /æG/ and /æN/ are generally consistent. For three speakers, /æN/ is even higher (lower F2) than the word list data; it is only lower for one speaker. Though there are insufficient data for /æG/, Speakers B and C show reliable pronunciations. Together with /æ/, these results are consistent with the /æG/- and /æN/-raising pattern seen in 3.3.4.

The IPI results appear to be the most reliable: /u:/ does not differ between styles at all and /æ/ does not differ in F2. This means the IPI values (i.e., mean F2 of /æ/ minus mean F2 of /u:/) will be approximately the same for these speakers. Though these speakers are only from Niagara, ON, this preliminarily supports the high innovation for Canadians seen in 3.3.5.

For Canadian Raising, the differences are more individual. Speaker A shows less fronting of /auT/ though height remains the same. Speaker B has a higher and fronter /ai/ but no change in /aiT/, lessening the degree of raising. For Speaker C, /ai/ becomes closer to /aiT/ in F2 but /aiT/ is raised more. Additionally, /au/ is further back, likely increasing the degree of /au/-raising. Speakers D and E have higher /aiT/, though D also fronts /au/ more. Speaker F has lower F1 for both /au/ and /auT/, such that the relative degree of raising should remain the same but both vowels will be higher. These varying results suggest that Canadian Raising is more influenced by style; more speakers would need to be analyzed to compare the actual degree of raising in both styles and to determine how much individual variation affects these results.

Finally, foreign (a) appears to be relatively unaffected by style. However, this sample is not representative: *pasta* was the only foreign (a) variable that could be identified outside of directly discussing language (i.e., when participants were asked directly about *pasta* and other foreign (a) words). Thus, the measurements for *pasta* are largely consistent between styles with the only difference being increased height (Speaker E). This supports the finding that Niagara, ON speakers tend to use the back (a) pronunciation for *pasta*. Further spontaneous data analysis would reveal if this is true across other regions and speakers, and may allow me to identify more foreign (a) variables in spontaneous speech for comparison.

3.3.9 Qualitative results

In contrast with the Likert scales in the questionnaire, the acoustic study assessed attitudinal information in the interviews with various questions about Canada and the United States. For example, participants were asked how similar they thought Canadians and Americans were, how similar Canadians and Americans sounds, and what specific things make someone sound Canadian or American to them. While these questions lend themselves to a qualitative rather than quantitative, correlational analysis, they also provide more specific information.

The question of how similar Canadians and Americans are generally depended on the participant's interpretation of what aspects the question covered. For example, many participants considered the people themselves to be very similar but political aspects to be very different.

Similarly, Canadians generally viewed Americans themselves positively and described positive encounters with individual people but viewed the more general culture and politics negatively in comparison with Canada. There seemed to be a split between those who mentioned this divide in their positive and negative opinions and those who only mentioned negative aspects, though these still tended to focus on political views rather than opinions on people. On the other side, Americans tended to report positive opinions about both Canadians and Canada.

In terms of speech, participants tended to agree that Americans (in Niagara, NY) and Canadians mostly sound similar but that they could be identified by some vocabulary and pronunciation differences. I asked for examples of pronunciations that make someone sound "American" or "Canadian." The Americans most often mentioned the pronunciation of *about* (Canadian Raising) and *sorry*. The Canadians most often identified aspects of the Northern Cities Shift, imitating fronted LOT (e.g., *"pop"*) and tensed TRAP (e.g., *"cat"*) vowels. There was generally a negative attitude associated with both features. They also tended to exaggerate the fronting of LOT such that it sounded more like their own /æ/ or even tensed /æ/.

Most interestingly, this misinterpretation of LOT-fronting also affected their perception of the foreign (a) variables. I specifically asked about the *pasta* variable since it showed the largest split in the questionnaire: if the participant used the back (a) in *pasta*, I asked them who they think says /pæstə/ and vice versa. When Canadians with back (a) were asked this question, they almost unanimously identified the front (a) variant as American. For example, the following quotes all come from Niagara, ON participants who used back (a) themselves:

"It would be /pæ/- /pjæstə/ more in the States, right?"
"/pæstə/ seems like more of an American thing because I feel like their A's come out in more like an [æ̃]."
"/pæstə/ is American for sure."

Furthermore, when imitating the front (a) variant they tended to exaggerate it to tensed /æ/. This also extended to other foreign (a) variables for some participants, particularly *taco*. While the proximity of the American fronted LOT vowel and the Canadian retracted TRAP vowel could account for the assumption that the Americans use /æ/, it does not explain why the Canadians usually imitated tensed /æ/ in these cases. Additionally, these (misperceived) pronunciations and tensed /æ/ more generally were typically stigmatized by Canadians.

3.4 Analysis

As with the questionnaire, we primarily see the border acting as a barrier for diffusion. This is true in both directions with the expected differences largely maintained. As predicted, the Americans clearly show the Northern Cities Shift pattern for /ɑ/, /ɔ/, and /æ/ while the Canadians instead merge /ɑ/-/ɔ/ and show the Low-Back-Merger-Shift pattern. The raising of the pre-nasal and pre-velar /æ/ allophones also follows the expected patterns for New York and Ontario. Canadians additionally show greater /u:/-fronting and /æ/-retraction, leading to more innovative IPI scores, as has been previously reported (Boberg 2010, 2021). Finally, the Canadians follow the expected Canadian Raising patterns for /au/ and /ai/ and the Americans also show strong /ai/-raising, which was expected in New York and is not indicative of diffusion.

However, several results do show potential diffusion and change. First, the foreign (a) results show an interesting pattern which expands on the questionnaire results. We again see a clear difference between Niagara, ON and the GTA with Niagara using more of the back (a) variant, particularly for *pasta*. The backing of (a) also only shows an increase in apparent time in Niagara, ON. This suggests wavelike diffusion of foreign (a) pronunciation. However, the acoustic analysis also makes it clear that Niagara, ON is not using the same vowel quality as Niagara, NY. Both Canadian regions tend to use their back /ɑ/-/ɔ/ vowel or front /æ/ vowel, with Niagara favouring /ɑ/-/ɔ/ more than Toronto. On the other hand, Niagara, NY tends to use their fronted /ɑ/, along with their raised and fronted /æ/ in particular words. This favours a phoneme-matching account where Canadians match (a) to either front /æ/ or back /ɑ/-/ɔ/, unlike Boberg's (2009) finding that central pronunciations may be used to more closely match the fronted /ɑ/ variant. However, the data here suggest intermediate pronunciations do sometimes occur, seemingly dependent on the word and speaker.

Importantly, there is also a clear divide in which words tend to have more front or back (a) and which are more variable. *Iran* and *pajamas* favour front (a) in all three regions and are raised and fronted like the normal /æN/ tokens are. The Canadian regions additionally favour front (a) for *Iraq*, *bratwurst*, and *panache*. On the other hand, *avocado*, *drama*, *llama*, *façade*, *taco*, and more overwhelmingly favour back (a) in all regions. In general, more words align with /ɑ/ or /ɑ/-/ɔ/ in all three regions.

Though the apparent time correlations only show increased backing with year of birth for Niagara, ON, these results overall show that foreign (a) continues to move towards the back

pronunciation in Canada: many more words favour back (a) in both regions than 23 years ago (see Boberg 2000). A larger sample in the GTA may reveal a similar correlation. It is however not clear if this is diffusion from the border or a more general "global prestige" associated with the back (a) variant, considering the actual quality varies on either side of the border. Additionally, the misperception of front (a) in *pasta* and *taco* as American and the stigmatization of the American LOT and TRAP vowels reported by many Canadians may suggest Canadians are simultaneously moving towards the global prestige variant of back (a) (corresponding to merged LOT-THOUGHT) *and* away from the stigmatized fronted variant (corresponding to unmerged LOT). In particular, this may drive the greater difference in *pasta* near the border since the Niagara, ON speakers will have greater exposure to the stigmatized Niagara, NY variant which they are misattributing to front (a).

For Canadian Raising, the Americans seem to also have /au/-raising. Though this raising is less strong than their /ai/-raising or the Canadians pattern, it is still significant and meets the threshold for Canadian Raising set out in the Atlas of North American English (Labov et al. 2016). The lack of any correlation with year of birth suggests this is not an emerging pattern, though it is then not decreasing either. Since there are only 12 speakers in Niagara, NY, this pattern requires further investigation to see if the threshold for raising is consistently met or if this is only the phonetic consequence of vowel shortening before voiceless obstruents. The spontaneous speech results suggest Canadian Raising may be affected by style and individual variation; a comparison between styles for Niagara, NY speakers may reveal a clearer pattern in either direction. The word list in this study also only focuses on environments for Canadian Raising; it is possible that the /au/- and /ai/-raising are a more generalized form of raising (i.e., not limited to the voiced-voiceless obstruent contrast) (Chambers 1973). However, if this proves to be phonological raising then the question remains whether it is a result of diffusion from across the border or elsewhere. Considering the New York raising is only a height increase without fronting, it could be phonologization of the phonetic shortening effect rather than an adopted feature, as Carmichael (2020) suggests for the emergence of /au/-raising in New Orleans.

At the same time, we see patterns of potential change in Canadian Raising for Canada. In Niagara, ON, there is an increase in fronting (higher F2 values) of /auT/ for younger speakers. This seems to indicate strengthening of raising, particularly in relation to Niagara, NY: as Americans only raise in height, the increased use of fronting distances the Canadians further from the

American pattern. This finding may then be an example of Canadian speakers maintaining and enhancing their local speech pattern, similar to Burnett's (2016) findings of increased Canadian variants closer to the border. However, there is no such correlation for /aiT/. It is therefore possible that fronting is instead being enhanced in relation to other regions of Canada which do not front /auT/ (Boberg 2010). This does not however explain the difference between Niagara and Toronto: there is no significant apparent time correlation in the GTA, but the borderline results for the height of /auT/ and /aiT/ suggest a potential decrease in the degree of raising. This pattern would require further investigation with a larger sample to see whether there is any correlation. It should be noted that no change was found for fronting which is an important part of raising in this region and the tokens are still raised in comparison to /au/ and /ai/, meaning any difference is only in degree or quality and not in the Canadian Raising pattern. In comparison with the Niagara, ON fronting pattern, raising in the GTA therefore seems stable. The potential fronting increase in Niagara may thus reflect a border effect since the GTA does not also show this enhancement.

The results for *houses* and *mouths* also present an interesting case of potential change in Canadian Raising. In both Canadian regions, *houses* is frequently raised despite the voiced /z/. In Niagara, ON, *mouths* is also raised for many speakers. While this pattern does not seem to occur across all voiced obstruent contexts and was not identified for /ai/, it suggests possible generalization of the raised vowel beyond the expected context (i.e., before voiceless obstruents). One possibility is that this is limited to plural versions of nouns, such that speakers interpret the raised /au/ in *house* and *mouth* as the phoneme and generalize it to the plural form. The pattern was only identified in the two singular-plural /au/ pairs in the word list, lending support to this idea; however, a more extensive word list focused on Canadian Raising could show whether this raising occurs elsewhere or may be word-specific. More data from the GTA would also show whether a portion of speakers do have *mouths*-raising like in Niagara.

Another interesting pattern that emerges from the acoustic results is the apparent time correlations for the low back merger. Niagara, NY shows a clear increase in merging for younger speakers. Though none of these speakers reach full merger, many do overlap with Canadian speakers. In conjunction with retraction of /æ/ for younger speakers, this may indicate reversal of the NCS pattern as has been reported in some areas (e.g., Driscoll & Lape 2015). As with the Canadian Raising pattern in Niagara, NY, it is not clear if this development is influenced by the border. However, these speakers clearly have contact with many Canadian speakers with the

merger and retraction which may enhance this pattern, even if it comes primarily from community-internal developments.

At the same time, the border again acts like a barrier even for these cases of possible diffusion: back (a) is still much more consistent for Americans than Canadians, /au/-raising is weak in New York, and the NCS patterns still largely remain intact even if the correlations show some decline. We also clearly see less potential cross-border diffusion in this study, with most of it linked to foreign (a); this is as expected given that structural variables should be less likely to diffuse (Labov 2007). As with the questionnaire, the border therefore seems to have a weakening effect on diffusion when it does occur in addition to preventing diffusion of many variables. We thus again see that the border plays a crucial role in diffusion and non-diffusion of pronunciation.

4. General discussion

Both the questionnaire and acoustic data have shown that the effect of the border on diffusion varies. Overall, Canadians remain distinct from Americans: only a handful of the questionnaire variables showed non-significant differences and the expected phonetic and phonological differences in the acoustic study were clearly demonstrated. Even when Canadians showed trends in the direction of American usage, the difference tended to remain significant. However, while the border most often acts as a barrier separating Canadian and American English, there is also evidence for both hierarchical and wave diffusion to some extent.

As stated, most variables showed little evidence of diffusion. All of the spelling variables and nearly all of the vocabulary variables favoured the Canadian variant in both Canadian regions. Both regions also strongly differed from Niagara, NY and did not differ from each other, suggesting no wavelike or hierarchical pattern of diffusion. Though there are small correlations which may suggest an increase in American forms, both regions remain overwhelmingly Canadian. This therefore suggests speakers are maintaining the Canadian English forms and resisting American influence, since the proximity between speakers would otherwise predict diffusion. This is in line with previous findings, including Easson's (1999) study of the Dialect Topography data in Niagara: most forms show a sharp drop-off at the border, maintaining the divide between Canadian and American English.

The biggest case for diffusion is the questionnaire pronunciation data where Niagara, ON and the GTA showed different frequencies for Canadian and American usage. An aggregate analysis of these frequencies found that pronunciation in Niagara, ON is generally more American and less Canadian than in Toronto. However, individual variable analyses revealed that only three variables have more frequent American usage in Niagara, ON: *pasta*, *caramel*, and *niche*. These variables reflect a pattern of wave diffusion where American variants are diffusing to nearby Niagara before reaching Toronto. In contrast, *asphalt* and *route* both showed hierarchical diffusion as Toronto had greater frequencies of the American variants than Niagara. Thus, the pattern of diffusion largely depends on the specific variable being studied. In particular, the very large difference for *pasta* suggests diffusion can be word-specific.

The foreign (a) variables like *pasta* are also more generally a strong case for diffusion. While *pasta* shows a divide, the other four variables in the questionnaire highly favour the

American back (a) pronunciation in both Canadian regions. Though the frequencies still significantly differed from Niagara, NY, all three regions clearly prefer back (a) in *drama*, *llama*, *taco*, and *avocado*. This is further confirmed in the acoustic study: the means for many of the words, including the four questionnaire variables, align more with the back vowel in all regions. We also again see a wavelike pattern with Niagara, ON ahead of the GTA for some words. In particular, *pasta* again shows a clear split between front (a) in Toronto and back (a) in Niagara. Niagara, ON additionally shows apparent time change with (a) further back among younger speakers overall and for *pasta* specifically.

However, the acoustic data also revealed that the (a) pronunciations in Niagara, NY are not the same as the Canadian pronunciations: front (a) corresponds to the tensed /æ/ for Americans and retracted /æ/ for Canadians (excluding raised /æN/), while back (a) corresponds to fronted, unmerged /ɑ/ for Americans and back, merged /ɑ/-/ɔ/ for Canadians. The increasing adoption of back (a) in Canada thus does not necessarily indicate diffusion from across the border. In the questionnaire, the higher use of back (a) in *pasta* specifically suggested wavelike diffusion from Niagara, NY to Niagara, ON. Though the acoustic results confirm that back (a) in *pasta* is much more common in Niagara, ON than the GTA, the different vowel qualities make it unclear whether this is actually an example of wavelike diffusion. These findings clearly demonstrate the shortcomings of the questionnaire: while both Niagara, NY and Niagara, ON overwhelmingly chose the same answer choice ("possible") for *pasta*, the actual production of the (a) varies between the two regions just as the Canadian /ɑ/-/ɔ/ and American /ɑ/ in *possible* vary.

It is possible that the Niagara, ON speakers are adopting the Niagara, NY variant but adjust it to their own back phoneme. However, the qualitative interview results show that the Canadians perceive the American back (a) as fronted, even linking it to the NY raised /æ/. An alternative possibility then is that the increased use of /ɑ/-/ɔ/ is a result of general diffusion from American media rather than the border; in that case, the use of back (a) would be associated with more global prestige. It could also be a response to the (mis)association of /pæstə/ with NY, given the negative attitudes the Niagara, ON participants tended to report for raised, fronted /æ/; in this case, the Niagara, ON speakers would be further distinguishing themselves from the "negative" pattern by using back /ɑ/-/ɔ/ and thus strengthening the barrier at the border. These two explanations need not be contradictory: if Canadians are generally moving towards more back (a) pronunciations as previously reported (Boberg 2000, 2009, 2020), this change may be accelerated in Niagara, ON in

response to the fronted variant in Niagara, NY. The pronunciation would therefore be increasingly similar to the general American pronunciation, but increasingly different from the Inland North pronunciation. This account supports Bailey et al.'s (1993) argument for the role of social and attitudinal factors in diffusion, including internal versus external prestige of specific variables, as the use of /ɑ/-/ɔ/ in Canada can be attributed to both global prestige and negative attitudes towards the NY variant.

Overall, pronunciation is therefore the most likely type of feature to diffuse. While some of this diffusion could be through media rather than the border, we clearly see the most similarities between Americans and Canadians for pronunciation. This is mostly limited to non-structural features like the questionnaire variables and foreign (a). These types of pronunciation variables deal with phonemic incidence, meaning speakers only need to learn which variants use a given phoneme within their system. On the other hand, the phonetic and phonological variables tended not to show diffusion. These variables would involve structural change in the speaker's system, making them less likely to be learnable for adult speakers (Labov 2007). This is also in line with Boberg's (2000) findings that foreign (a) does diffuse across the Detroit-Windsor border while differences in speakers' phonemic systems do not. While /au/-raising, the low back merger, and /æ/-retraction all show some potential use in New York, the effects are much smaller. Additionally, these may be community-internal developments (e.g., the phonemicization of phonetic shortening for /au/-raising) rather than diffusion from Canada; they may however be enhanced through contact with Canadian English speakers.

Importantly, these instances of apparent diffusion still support a strong border effect. If diffusion were simply a matter of size and distance, both Niagara, ON and the GTA should be much more similar to Niagara, NY. Even if Canadians are adopting American foreign (a) and certain other pronunciations, this change is clearly slowed down by the border. The border therefore has a weakening effect on diffusion, whether this means reducing the degree of diffusion or blocking it entirely. This supports Boberg's (2000) concerns about Trudgill's (1974) gravity model: the model must be modified to account for wave diffusion and the varying effects of a national border. While these varying patterns may be difficult to actually quantify as one "border effect," it is clear that we cannot account for the diffusion and non-diffusion we see in Niagara simply in terms of distance and size.

We also see that the border effect may be driven by attitudes and the maintenance of identity, as previously reported (e.g., Allen 1959, Miller 1989, Easson 1999, Burnett 2016, Swan 2020). In both the questionnaire ideological scales and the interview questions, Canadians generally expressed negative attitudes towards American English, particularly that of Niagara, NY. The interviews further revealed that this was largely related to the Northern Cities Shift and the fronted LOT and tensed TRAP vowels in particular. Though their assessment was not always accurate, as described above for foreign (a), these features were apparently quite salient and stigmatized. Increasing social salience of the NCS and stigmatization by younger speakers has been linked to the possible reversal of NCS patterns (Driscoll & Lape 2015). These attitudes may therefore account for the Canadians' resistance to pronunciations perceived as similar to fronted LOT/tensed TRAP (e.g., the back (a) pronunciation in Niagara, NY). Additionally, though the NCS features were not frequently mentioned by the Americans in the interviews, this stigmatization may explain the patterns of NCS reversal seen here (e.g., retraction of /æ/ for younger speakers, more merging of /ɑ/-/ɔ/ for younger speakers).

Social salience also seems to affect which questionnaire variables showed more or less diffusion. Though participants were not specifically asked about these variables, many mentioned vocabulary as identifiably Canadian or American in the interviews. When pronunciation was mentioned, it focused more on structural differences like Canadian Raising and the NCS rather than the phonemic incidence type of variables studied in the questionnaire. If vocabulary is indeed more salient as a Canadian/American marker, the maintenance of vocabulary differences may be related to the maintenance of local identity. This again supports previous findings regarding attitudinal and identity factors (e.g., Bailey et al. 1993, Miller 1989, Burnett 2006, Swan 2020): diffusion is variable-dependent and likely influenced by the social evaluation and association of the possible variants.

Beyond diffusion, the findings also reveal more general patterns of stability and potential change in Canadian English. The Low-Back-Merger Shift and raising of /æG/ and /æN/ are clearly demonstrated for the Canadians. Though both show no strengthening in apparent time, they also show no decline, indicating stability of these Canadian features. Canadian Raising also remains largely stable. The Niagara, ON speakers do show increased fronting of /auT/, suggesting potential strengthening of the contrast. We also see generalization of the raised pronunciation in *houses* and

possibly *mouths*, providing initial evidence for use of the raised diphthongs beyond the phonological conditioning context.

There are however still some gaps and remaining questions. First, the aggregate analysis of the questionnaire largely depended on being able to categorize variants as "Canadian" or "American." This meant unclear variants could not be coded and were left out; while this makes sense for focusing on identifiable Canadian-American differences, it also means some information about the Niagara and Toronto patterns is lost. This is also true for the acoustic study as only the major expected Canadian-American differences were analyzed. The large amount of data will allow for more in-depth future analyses. Additionally, the analysis here has been limited to basic individual comparisons via ANOVAs; the use of regression models would allow for greater comparison of effects and will be an important part of further work on the effect of the border. It should also be noted that the acoustic data here come primarily from word list speech; though the preliminary comparison with spontaneous speech predicts that the findings will mostly hold, a word list is of course not natural speech and cannot fully capture the effect of diffusion in everyday language.

Another concern is that we cannot distinguish cross-border influence from the more general influence of exposure to American people and media. The former is particularly important for somewhere like Niagara where tourists come from everywhere, not just immediately across the border. On both sides, there are greater opportunities for contact with speakers from many dialect areas than in a more typical border town. The effect of media on diffusion is also important and cannot be adequately separated from other influences (Boberg 2021). In particular, I have suggested that the adoption of American foreign (a) pronunciations may be affected by both exposure to the American pronunciation in media and contact with the more fronted variant in Niagara, NY. Though we can see the continuous diffusion of this variable and the difference between Niagara and Toronto, we cannot separate the different sources of diffusion.

Similarly, this study cannot identify whether apparent diffusion in Niagara, NY is from Canada or elsewhere. While Niagara, ON has the GTA as a comparison region, there is no comparison region in New York. This decision was partially due to practicality of recruiting and interviewing enough participants from a fourth region and partially due to the assumption that American English has a greater effect on Canadian English than vice versa. While previous studies have reported greater adoption of American features in Canada (e.g., Avis 1954, 1955, 1956; Allen

1959), at least some diffusion is typically reported in both directions. To see if Niagara, NY is affected by cross-border diffusion, especially for features like /au/-raising and NCS reversal identified here, future research will need to include an additional comparison region.

Returning to the main research questions, the studies presented here make it clear that the border has an effect on diffusion and that diffusion is much more complex than can be accounted for by size and distance. The national border has a general dampening effect on diffusion and most frequently acts as a linguistic barrier. When there is diffusion, the pattern also varies between wavelike diffusion with Niagara, ON adopting American variants more quickly and hierarchical diffusion with the GTA adopting American variants more quickly. In many cases, both Canadian regions also showed equal rates of diffusion or non-diffusion and therefore did not follow either pattern clearly. Additionally, some pronunciation features like Canadian Raising, the low back merger, and retraction of /æ/ showed a possible influence of Canadian English on the American speakers, though it is unclear if this is a border effect. In either case, the type of feature and the specific variable clearly affected the pattern of diffusion with non-structural pronunciation features showing greater potential for diffusion, in line with previous findings (e.g., Boberg 2000, Labov 2007). The social evaluation and salience of different types of features also seems to have an effect, suggesting identity plays a role in diffusion.

5. Summary and conclusion

This thesis has examined the role of the national border on linguistic diffusion in Niagara. The primary goal was to determine whether the border has an effect by looking at patterns of diffusion and non-diffusion for various Canadian-American differences in pronunciation, vocabulary, grammar, and spelling. Overall, the results have shown that the border is mainly a barrier to diffusion. If the border had no effect, we would expect a greater amount of wavelike diffusion to Niagara, ON due to proximity and of hierarchical diffusion to the GTA due to size. Instead, we see that Canadian and American English largely remain separate. Though some diffusion does occur, it is weakened relative to what would be predicted based solely on distance and size.

I also specifically compared the different types of variables to see which were more likely to diffuse or not. In both the questionnaire and acoustic study, pronunciation of phonemic incidence variables like foreign (a) showed greater diffusion. Though pronunciation generally followed a wavelike pattern with Niagara, ON as an intermediate dialect area, specific variables showed both wavelike and hierarchical patterns in addition to non-diffusion. In particular, the pronunciation of *pasta* showed a large split between Niagara and Toronto. On the other hand, structural pronunciation features, vocabulary, and spelling showed much less propensity for diffusion. The grammar questions also highlighted the effect of individual variables, with older variables showing no difference but *done (with)* showing a large Canadian-American divide.

These findings go against a common assumption that Canadians near the border sound more "American." Many participants expressed this sentiment themselves or recounted stories of being confused for Americans in Canada. Though some of the phonemic incidence variables do show diffusion, the major American pronunciation features like the NCS do not and the major Canadian features like Canadian Raising and the LBMS appear to be unaffected. In most cases, Niagara, ON speakers seem to sound no more "American" than the Toronto speakers.

The border effect and varying patterns of diffusion seem to be affected by the social salience of the variables. In particular, vocabulary seems to be particularly identifiable as "Canadian" and the preservation of vocabulary differences may therefore reflect speakers maintaining their Canadian identity. Meanwhile, salient features that are viewed negatively like the NCS tensed TRAP and fronted LOT vowels may lead to resistance of variants perceived as

similar; specifically, this may account for the increased use of back (a) with the far back LOT-THOUGHT vowel in Niagara, ON (e.g., in *pasta*). These social evaluation effects could also explain the patterns of stability and change: while Canadian features like the LBMS and Canadian Raising are largely stable in Canada, the NCS shows potential reversal in Niagara, NY.

Overall then, the border has an effect on diffusion which depends on the specific features and their social evaluation. It is therefore clear that diffusion cannot be captured solely through size and distance like Trudgill's (1974) model would suggest. Additionally, the effects of size and distance vary as the results here demonstrate instances of both wavelike and hierarchical diffusion. While the gravity model provides a starting point for expected diffusion, it would need many more factors to predict the patterns seen here.

More generally, these findings show the interaction between identity and linguistic diffusion. Canadians often want to set themselves apart from Americans and language is one way to do this. The maintenance of distinctive Canadian features like vocabulary, spelling, Canadian Raising, and the Low-Back-Merger Shift allows speakers to reassert their Canadian identity. These features additionally help speakers differentiate themselves from the negatively perceived American patterns like the Northern Cities Shift. On the other hand, some American patterns are evidently viewed positively such as back (a) pronunciations in foreign loanwords. In some cases, these judgments conflict such as the fronted back (a) pronunciation in New York, seemingly leading to faster adoption of the prestigious variant. In Niagara, these complex interactions appear to be enhanced by the border and the increased contact between speakers. The study of diffusion therefore must consider not only the effect of national borders but how the variables of interest interact with national identity.

References

Allen, Harold B. 1959. Canadian-American speech differences along the middle border. *Canadian Journal of Linguistics / Revue Canadienne de linguistique* 5(1). 17-24.

Avis, Walter S. 1954. Speech differences along the Ontario-United States border. *Canadian Journal of Linguistics / Revue Canadienne de linguistique* 1(1).13-18.

Avis, Walter S. 1955. Speech differences along the Ontario-United States border. *Canadian Journal of Linguistics / Revue Canadienne de linguistique* 1(R1). 14-19.

Avis, Walter S. 1956. Speech differences along the Ontario-United States border. *Canadian Journal of Linguistics / Revue Canadienne de linguistique* 2(2). 41-59.

Avis, Walter S. 1973. The English language in Canada. In T. A. Sebeok (ed.), *Linguistics in North America* (Current Trends in Linguistics 10), 40-74. The Hague: Mouton.

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5. 359-390.

Bloomfield, Leonard. (1933). *Language*. New York: Holt.

Boberg, Charles. 2000. Geolinguistic diffusion and the U.S.-Canada border. *Language Variation and Change* 12. 1-24.

Boberg, Charles. 2005. The North American Regional Vocabulary Survey: Renewing the study of lexical variation in North American English. *American Speech* 80(1). 22-60.

Boberg, Charles. 2008. Regional phonetic differentiation in Standard Canadian English. *Journal of English Linguistics* 36(2). 103-188.

Boberg, Charles. 2009. The emergence of a new phoneme: Foreign (a) in Canadian English. *Language Variation and Change* 21(3). 355-380.

Boberg, Charles. 2010. *The English language in Canada: Status, history and comparative analysis*. Cambridge: Cambridge University Press.

Boberg, Charles. 2019a. A closer look at the Short Front Vowel Shift in Canada. *Journal of English Linguistics* 47(2). 91-119.

Boberg, Charles. 2019b. The North American Low-Back-Merger Shift: A continental sound change. In Becker, Kara (ed.), The Low-Back-Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and Short Front Vowel Shifts across North America (Publication of the American Dialect Society 104), 56-73. Durham, NC: Duke University Press.

Boberg, Charles. 2020. Foreign (a) in North American English: Variation and change in loan phonology. *Journal of English Linguistics* 48(1). 31-71.

Boberg, Charles. 2021. *Accent in North American film and television: A sociophonetic analysis.* Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink. 2021. Praat: Doing phonetics by computer. Version 6.2.01. http://www.praat.org (November 17, 2021.)

Burnett, Wendy. 2006. Linguistic resistance on the Maine-New Brunswick border. *Canadian Journal of Linguistics / Revue Canadienne de linguistique* 51(2-3). 161-176.

Callary, Robert E. 1975. Phonological change and the development of an urban dialect in Illinois. *Language in Society* 4. 155-169.

Cambridge University Press. 2023. Niche. https://dictionary.cambridge.org/pronunciation/english/niche (10 June, 2023).

Carmichael, Katie. 2020. The rise of Canadian raising of /au/ in New Orleans English. *The Journal of the Acoustical Society of America* 147(1). 554-567.

Chambers, J. K. 1973. Canadian raising. *Canadian Journal of Linguistics / Revue Canadienne de linguistique* 18. 113-135.

Chambers, J. K. 1990. The Canada-US border as a vanishing isogloss: The evidence of *chesterfield*. *Journal of English Linguistics* 23(1-2). 155–166.

Chambers, J. K. 1994. An introduction to dialect topography. *English World-Wide* 15. 35–53.

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

Driscoll, Anna & Emma Lape. 2015. Reversal of the Northern Cities Shift in Syracuse, New York. *University of Pennsylvania Working Papers in Linguistics* 21(2). 41-47.

Easson, Gordon. J. 1999. Sounds of the Golden Horseshoe: Canadian-American differences at the Niagara border. *Toronto Working Papers in Linguistics* 17.

Hock, Hans H. 2021. *Principles of Historical Linguistics* (Trends in Linguistics Studies and Monographs 34). Berlin: Mouton de Gruyter.

Jankowski, Bridget L. & Sali A. Tagliamonte. 2019. Supper or dinner?: Sociolinguistic variation in the meals of the day. *English World-Wide* 40(2). 170-201.

Kurath, H. 1949. *A word geography of the eastern United States*. Ann Arbor: University of Michigan Press.

Labov, William. 1963. The social motivation of a sound change. *WORD* 19(3). 273-309.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1991. The three dialects of English. In P. Eckert (ed.), *New Ways of Analyzing Sound Change*, 1-44. New York: Academic Press.

Labov, William. 2007. Transmission and diffusion. *Language Variation and Change* 5. 359-390.

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

Merriam-Webster. 2023. Niche. https://www.merriam-webster.com/dictionary/niche (10 June, 2023).

Miller, Corey. 1989. The United States-Canadian border as a linguistic boundary: The English language in Calais, Maine and St. Stephen, New Brunswick. Cambridge, MA: Harvard University Bachelor of Arts honors essay.

Nearey, Terrance M. 1978. *Phonetic feature systems for vowels*. Bloomington, IN: Indiana University Linguistics Club.

Nerbonne, John. 2009. Data-driven dialectology. *Language and Linguistics Compass* 3(1). 175-198.

Nycz, Jennifer & Lauren Hall-Lew. 2013. Best practices in measuring vowel mergers. *Proceedings of Meetings on Acoustics* 20(1).

Scargill, M. H. & H. J. Warkentyne. 1972. The survey of Canadian English: A report. *The English Quarterly* 5(3). 47-104.

Swan, Julia T. 2020. *Bag* across the border: Sociocultural background, ideological stance, and *bag* raising in Seattle and Vancouver. *American Speech* 95(1). 46-81.

Trudgill, Peter. 1974. Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society* 3. 215-246.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Appendix

A. Questionnaire and response frequencies

Note: * indicates that the question was excluded from analyses due to difficulties coding variants. (C) indicates Canadian-coded variant and (A) indicates American-coded variant.

Note: Frequencies (%) are rounded to the nearest whole number. For vocabulary, only variants with at least 1% usage in at least one region are included; therefore, not all values add to 100%. Blank responses are not included in totals.

Vocabulary:

(1) When you've finished your meal at a restaurant, you ask for the _____.

Region	Bill (C)	Check (A)	Cheque
Niagara, ON	63	16	19
Greater Toronto Area	72	5	22
Niagara, NY	16	81	1

(2) When you're finished shopping, you pay at the _____.

Region	Cash (C)	Register (A)	Cash register	Till	Checkout
Niagara, ON	13	18	10	7	20
Greater Toronto Area	20	23	14	10	12
Niagara, NY	0	55	4	0	11

(3) When you want water from the sink or tub, you turn the _____.

Region	Tap (C)	Faucet (A)
Niagara, ON	84	11
Greater Toronto Area	87	11
Niagara, NY	15	74

(4) What do you call the room in your home with a toilet? _____.

Region	Washroom (C)	Bathroom (A)
Niagara, ON	12	87
Greater Toronto Area	21	73
Niagara, NY	1	95

(5) What do you call the room with a toilet in a public place? _____.

Region	Washroom (C)	Bathroom (A)	Restroom (A)
Niagara, ON	63	17	15
Greater Toronto Area	64	17	10
Niagara, NY	3	19	64

(6) What do you call your main evening meal? _____.

Region	Supper (C)	Dinner (A)
Niagara, ON	40	58
Greater Toronto Area	12	85
Niagara, NY	17	79

(7) What do you call a non-alcoholic carbonated drink (Ginger Ale, Pepsi, Sprite, etc.)? _____.

Region	Pop (C)	Soda (A)
Niagara, ON	93	4
Greater Toronto Area	92	3
Niagara, NY	72	21

(8) What do you call the PAPER used for wiping your hands or mouth while eating? _____.

Region	Serviette (C)	Napkin (A)	Paper towel
Niagara, ON	11	79	5
Greater Toronto Area	12	71	11
Niagara, NY	2	93	3

(9) What do you call the CLOTH used for wiping your hands or mouth while eating? _____.

Region	Serviette (C)	Napkin (A)	Cloth napkin
Niagara, ON	7	75	4
Greater Toronto Area	5	82	5
Niagara, NY	1	89	3

(10) What do you call the following symbols?

( ) : _____.

Region	Brackets (C)	Parentheses (A)
Niagara, ON	71	25
Greater Toronto Area	63	30
Niagara, NY	2	96

[ ] : _____.

Region	Square brackets (C)	Brackets (A)	Parentheses
Niagara, ON	40	38	5
Greater Toronto Area	57	23	4
Niagara, NY	2	61	10

* { }: _____.

Region	Curly brackets	Fancy brackets	Squiggly Brackets	Brackets	Parentheses	Braces
Niagara, ON	10	5	6	39	18	4
Greater Toronto Area	25	1	17	17	15	3
Niagara, NY	4	3	2	46	18	7

! : _____.

Region	Exclamation mark (C)	Exclamation point (A)
Niagara, ON	48	37
Greater Toronto Area	54	39
Niagara, NY	11	67

(11) What do you call the thick, fluffy blanket that keeps you warm in bed?: *duvet / comforter / quilt / other:* _____.

Region	Duvet (C)	Comforter (A)	Quilt
Niagara, ON	25	61	5
Greater Toronto Area	42	41	3
Niagara, NY	4	82	4

(12) What do you call the thick, sweet spread used to cover a cake?: *frosting / icing / other:* _____.

Region	Icing (C)	Frosting (A)
Niagara, ON	82	17
Greater Toronto Area	76	21
Niagara, NY	18	81

(13) What do you call the metal structure around the edge of the roof that collects rainwater?: *gutters / eavestroughs / other:* _____.

Region	Eavestroughs (C)	Gutters (A)
Niagara, ON	87	13

Greater Toronto Area	74	25
Niagara, NY	4	96

(14) What do you call the small cloth used for washing your face?: *washcloth / face cloth / other: _____*.

Region	Face cloth (C)	Washcloth (A)
Niagara, ON	51	49
Greater Toronto Area	61	38
Niagara, NY	5	94

(15) What do you call athletic shoes worn as casual attire?: *running shoes / sneakers / tennis shoes / runners / trainers / other: _____*.

Region	Running shoes (C)	Runners (C)	Sneakers (A)	Tennis shoes (A)	Trainers
Niagara, ON	68	17	12	0	0
Greater Toronto Area	58	10	25	1	2
Niagara, NY	2	0	94	2	0

* (16) Which of the following names do you use for a waterway smaller than a river? <u>You may select multiple options if you would normally use more than one word:</u> *crick / stream / brook / creek / other: _____*.

Region	Crick	Stream	Brook	Creek
Niagara, ON	26	59	6	80
Greater Toronto Area	0	68	14	92
Niagara, NY	21	52	18	85

* (17) When you didn't hear someone and want them to repeat themselves, you normally say (<u>You may select multiple options if you would normally use more than one word</u>): *Pardon (me)? / (I'm) sorry? / Excuse me? / What? / Huh? / other: _____*.

Region	Pardon (me)?	(I'm) sorry?	Excuse me?	What?	Huh?
Niagara, ON	66	35	14	19	10
Greater Toronto Area	70	49	24	20	10
Niagara, NY	37	24	31	31	11

Pronunciation:

(18) Do you pronounce the letter Z as zee or zed?

Region	Zed (C)	Zee (A)
Niagara, ON	73	27
Greater Toronto Area	78	22
Niagara, NY	5	95

(19) Does LEISURE rhyme with seizure or pleasure?

Region	Pleasure (C)	Seizure (A)
Niagara, ON	7	93
Greater Toronto Area	12	88
Niagara, NY	5	95

(20) In ASPHALT, does the AS sound like ash?

Region	Yes (C)	No (A)
Niagara, ON	89	11
Greater Toronto Area	79	21
Niagara, NY	38	62

(21) Does the first part of EITHER sound like the ie of pie or the ee of bee?

Region	Pie (C)	Bee (A)
Niagara, ON	29	71
Greater Toronto Area	30	83
Niagara, NY	17	70

(22) Does the O of PRODUCE (noun, as in fresh produce from the store) sound like the o of go or the o of got?

Region	Got (C)	Go (A)
Niagara, ON	19	81
Greater Toronto Area	30	70
Niagara, NY	1	99

(23) Does the first part of PASTA sound like the first part of passing or possible?

Region	Passing (C)	Possible (A)
Niagara, ON	42	58
Greater Toronto Area	73	28
Niagara, NY	12	88

(24) Does DRAMA rhyme more with comma or gamma?

Region	Gamma (C)	Comma (A)
Niagara, ON	18	82
Greater Toronto Area	20	80
Niagara, NY	7	93

(25) Does the first part of TACO sound more like tack or tock?

Region	Tack (C)	Tock (A)
Niagara, ON	8	92
Greater Toronto Area	10	90
Niagara, NY	3	97

(26) Is CARAMEL pronounced with two or three syllables?

Region	Two (A)	Three (C)
Niagara, ON	27	74
Greater Toronto Area	10	90
Niagara, NY	55	45

(27) Does the O of PROGRESS (noun, as in making progress) sound like the o of go or the o of got?

Region	Go (C)	Got (A)
Niagara, ON	44	56
Greater Toronto Area	43	81
Niagara, NY	19	57

(28) Does LLAMA rhyme more with comma or gamma?

Region	Gamma (C)	Comma (A)
Niagara, ON	27	73
Greater Toronto Area	23	77
Niagara, NY	8	92

(29) Does the cad in AVOCADO rhyme more with pod or pad?

Region	Pad (C)	Pod (A)
Niagara, ON	12	88
Greater Toronto Area	6	85
Niagara, NY	15	94

(30) Does AGAIN rhyme more with pane, pin, or pen?

Region	Pane (C)	Pen (A)	Pin
Niagara, ON	24	74	2
Greater Toronto Area	21	75	4
Niagara, NY	6	88	6

(31) Do you pronounce the L in ALMOND?

Region	Yes (A)	No (C)
Niagara, ON	62	38
Greater Toronto Area	62	38
Niagara, NY	74	26

(32) Does the O in PROJECT (noun) sound like the o in go or the o in got?

Region	Go (C)	Got (A)
Niagara, ON	30	70
Greater Toronto Area	24	76
Niagara, NY	15	85

* (33) Does PECAN sound more like PEE-can, PEE-cahn, or puh-CAHN? Capital letters indicate the main stress (the part that sounds louder).

Region	PEE-can	PEE-cahn	puh-CAHN
Niagara, ON	75	14	11
Greater Toronto Area	80	7	13
Niagara, NY	57	22	21

* (34) Does BAGEL sound more like BEG-uhl, BAG-uhl, or BAYG-uhl? Capital letters indicate the main stress (the part that sounds louder).

Region	BEG-uhl	BAG-uhl	BAYG-uhl
Niagara, ON	11	19	70
Greater Toronto Area	7	17	76
Niagara, NY	5	20	75

(35) Does NICHE rhyme with ditch or quiche?

Region	Ditch (A)	Quiche (C)
Niagara, ON	9	91
Greater Toronto Area	1	99
Niagara, NY	44	56

* (36) Does ROUTE rhyme with shout or shoot?

Region	Shout	Shoot
Niagara, ON	33	67
Greater Toronto Area	19	81
Niagara, NY	41	59

Grammar/phrases:

(37) Which do you say?: *Our house is very different to yours. / Our house is very different than yours. / Our house is very different from yours.*

Region	to	than (A)	from (C)
Niagara, ON	2	60	3
Greater Toronto Area	3	52	45
Niagara, NY	1	57	52

(38) Which do you say?: *He sneaked by when my back was turned. / He snuck by when my back was turned.*

Region	sneaked (C)	snuck (A)
Niagara, ON	6	94
Greater Toronto Area	7	93
Niagara, NY	8	92

(39) Which do you say?: *He loaned me a dollar. / He lent me a dollar. / He borrowed me a dollar.*

Region	loaned (C)	lent (A)	borrowed
Niagara, ON	24	75	1
Greater Toronto Area	46	54	0
Niagara, NY	23	77	0

(40) Which do you say?: *She dived in the lake. / She dove in the lake.*

Region	dived (C)	dove (A)
Niagara, ON	4	96
Greater Toronto Area	3	97
Niagara, NY	5	95

(41) Which do you say?: *I'm done my homework. / I'm done with my homework.*

Region	done (C)	done with (A)
Niagara, ON	81	19
Greater Toronto Area	79	21
Niagara, NY	3	97

Spelling:

(42) Which spelling do you use?: *grey / gray / either one*

Region	grey (C)	gray (A)	either one
Niagara, ON	54	11	35
Greater Toronto Area	52	10	38
Niagara, NY	15	36	49

(43) Which spelling do you use?: *colour / color / either one*

Region	colour (C)	color (A)	either one
Niagara, ON	89	6	5
Greater Toronto Area	93	2	5
Niagara, NY	2	93	5

(44) Which spelling do you use for the bank document used to transfer money?: *check / cheque / either one*

Region	cheque (C)	check (A)	either one
Niagara, ON	92	4	4
Greater Toronto Area	93	1	6
Niagara, NY	2	94	4

(45) Which spelling do you use?: *traveled / travelled / either one*

Region	travelled (A)	travelled (C)	either one
Niagara, ON	10	81	9
Greater Toronto Area	12	77	11
Niagara, NY	74	15	11

(46) Which spelling do you use?: *center / centre / either one*

Region	center (A)	centre (C)	either one
Niagara, ON	16	54	30
Greater Toronto Area	7	63	30
Niagara, NY	91	2	7

B. Vowel guide and word list

Phoneme/ allophone	Wells keyword	Word list variables
(a)	–	avocado, bratwurst, Colorado, drama, façade, Iran, Iraq, llama, nachos, pajamas, panache, pasta, taco
/æ/	TRAP-BATH	bad, bath, cast, pass, sack, staff, trap
/æG/	TRAP	bag, bang, gag, gang, tag, tank
/æN/	TRAP-BATH	ham, pan, sand, stamp
/ær/	TRAP	barrel, carrot, marry
/ai/	PRICE	eyes, guy, high, knives, rider, side, tie
/aiT/	PRICE	height, ice, knife, price, sight, tight, wife, writer
/au/	MOUTH	cow, houses, loud, mouths, vow
/auN/	MOUTH	found, town
/auT/	MOUTH	about, doubt, house, mouth, shout, south
/ɑ/	LOT-PALM	bomb, bother, calm, cot, father, lot, palm, sock, sod, spa, stock
/ɑl/	LOT-PALM	collar
/ɑr/	LOT	borrow, sorry, tomorrow
(ahr)	START	bar, car, card, park, start
/ei/	FACE	bay, face, page, plague, say, take, vague
/eir/	SQUARE	bare, fairy, Mary
/ɛ/	DRESS	bed, dead, deck, dress
/ɛG/	DRESS	beg, keg, leg
/ɛN/	DRESS	hem, pen
/ɛr/	DRESS	berry, ferry, merry
/i:/	FLEECE	bee, fleece, key, seek
/ɪ/	KIT	bid, kiss, kit, sick
/ɪN/	KIT	him, pin
/oi/	CHOICE	boy, choice, toy, void
/ou/	GOAT	code, go, goat, toe
/oul/	GOAT	coal, fold, pole
/our/	FORCE	door, force
/ɔ/	THOUGHT-CLOTH	caught, cloth, dog, log, paw, sawed, stalk, song, thought
/ɔl/	THOUGHT-CLOTH	caller
/ɔr/	NORTH	north
/ɚ/	NURSE	bird, her, nurse
/u:/	GOOSE	do, food, goose, two
/u:l/	GOOSE	cool, pool
/ʊ/	FOOT	foot, good, took

/ʌ/	STRUT	bus, cuff, duck

C. Mean formant values (normalized) by phoneme in each region

All values are rounded to the nearest whole number. Values are in normalized Hertz.

	Phoneme	Mean F1 (Hz)	Standard Deviation	Mean F2 (Hz)	Standard Deviation
Niagara, ON	(a)	814	101	1397	248
Niagara, NY		872	133	1516	216
Greater Toronto Area		847	117	1439	236
Niagara, ON	/æ/	910	91	1638	135
Niagara, NY		685	106	1943	163
Greater Toronto Area		944	99	1643	110
Niagara, ON	/æG/	710	108	2062	200
Niagara, NY		679	100	1989	174
Greater Toronto Area		760	125	2020	219
Niagara, ON	/æN/	634	98	2231	209
Niagara, NY		556	115	2189	182
Greater Toronto Area		649	103	2166	188
Niagara, ON	/ær/	569	71	2054	168
Niagara, NY		513	78	1979	186
Greater Toronto Area		571	85	1954	183
Niagara, ON	/ai/	867	90	1397	125
Niagara, NY		929	89	1499	102
Greater Toronto Area		891	100	1395	132
Niagara, ON	/aiT/	748	91	1609	160
Niagara, NY		774	112	1459	149
Greater Toronto Area		782	101	1599	167
Niagara, ON	/au/	883	97	1568	130
Niagara, NY		917	101	1478	116
Greater Toronto Area		909	97	1559	123
Niagara, ON	/auT/	748	88	1617	160
Niagara, NY		866	113	1480	107
Greater Toronto Area		782	102	1646	126
Niagara, ON	/auN/	842	120	1735	159
Niagara, NY		865	118	1541	122
Greater Toronto Area		883	137	1752	114
Niagara, ON		791	82	1168	112

Niagara, NY	/ɑ/	908	86	1380	129
Greater Toronto Area		789	84	1149	113
Niagara, ON	/ɑl/	739	72	1075	69
Niagara, NY		876	115	1312	150
Greater Toronto Area		747	72	1088	139
Niagara, ON	/ɑr/	646	124	1034	183
Niagara, NY		782	121	1329	172
Greater Toronto Area		597	108	989	182
Niagara, ON	(ahr)	747	81	1282	148
Niagara, NY		823	102	1411	190
Greater Toronto Area		742	80	1235	168
Niagara, ON	/ei/	547	75	2193	201
Niagara, NY		537	65	2069	173
Greater Toronto Area		542	84	2131	211
Niagara, ON	/eir/	550	68	2133	174
Niagara, NY		507	70	2072	127
Greater Toronto Area		552	54	2076	136
Niagara, ON	/ɛ/	702	81	1903	155
Niagara, NY		711	88	1791	114
Greater Toronto Area		720	100	1883	127
Niagara, ON	/ɛG/	644	66	2038	179
Niagara, NY		672	80	1864	132
Greater Toronto Area		651	89	2025	155
Niagara, ON	/ɛN/	753	92	1911	169
Niagara, NY		775	138	1838	147
Greater Toronto Area		739	134	1916	180
Niagara, ON	/ɛr/	544	71	2111	158
Niagara, NY		497	66	2029	140
Greater Toronto Area		561	60	2046	122
Niagara, ON	/i:/	368	58	2637	207
Niagara, NY		378	46	2546	146
Greater Toronto Area		357	59	2644	136
Niagara, ON	/ɪ/	550	60	2072	144
Niagara, NY		521	63	1966	118
Greater Toronto Area		550	58	2051	113
Niagara, ON	/ɪN/	561	79	2162	191
Niagara, NY		506	91	2086	157
Greater Toronto Area		540	74	2140	180

Niagara, ON	/oi/	511	69	901	148
Niagara, NY		486	60	865	136
Greater Toronto Area		499	63	879	151
Niagara, ON	/ou/	593	58	1241	141
Niagara, NY		599	76	1190	182
Greater Toronto Area		603	69	1324	139
Niagara, ON	/oul/	513	53	824	97
Niagara, NY		529	64	848	115
Greater Toronto Area		512	62	824	98
Niagara, ON	/our/	491	58	800	111
Niagara, NY		456	57	780	107
Greater Toronto Area		485	64	814	127
Niagara, ON	/ɔ/	777	81	1167	107
Niagara, NY		801	106	1177	119
Greater Toronto Area		780	90	1140	110
Niagara, ON	/ɔl/	742	75	1078	84
Niagara, NY		757	67	1141	111
Greater Toronto Area		752	58	1081	97
Niagara, ON	/ɔr/	581	83	975	153
Niagara, NY		542	73	908	128
Greater Toronto Area		592	83	974	153
Niagara, ON	/ɚ/	551	64	1531	185
Niagara, NY		537	72	1423	221
Greater Toronto Area		567	57	1519	216
Niagara, ON	/u:/	393	51	1793	324
Niagara, NY		392	41	1447	315
Greater Toronto Area		393	60	1849	276
Niagara, ON	/u:l/	396	57	850	127
Niagara, NY		408	40	772	111
Greater Toronto Area		401	59	798	168
Niagara, ON	/ʊ/	579	77	1460	159
Niagara, NY		543	71	1323	155
Greater Toronto Area		566	74	1435	150
Niagara, ON	/ʌ/	792	76	1444	112
Niagara, NY		793	78	1350	117
Greater Toronto Area		804	77	1448	101

D. Mean formant values (normalized) by individual foreign (a) variable in each region

All values are rounded to the nearest whole number. Values are in normalized Hertz.

	Niagara, ON			
Word	**F1 (Hz)**	**Std. Dev.**	**F2 (Hz)**	**Std. Dev.**
avocado	825	77	1299	128
bratwurst	866	86	1519	144
Colorado	815	85	1338	150
drama	775	97	1250	155
façade	803	81	1247	122
Iran	724	114	1740	370
Iraq	869	85	1560	139
llama	794	84	1205	143
nachos	851	109	1360	125
pajamas	759	107	1605	293
panache	854	123	1462	228
pasta	839	70	1308	180
taco	817	81	1271	145

	Greater Toronto Area			
Word	**F1 (Hz)**	**Std. Dev.**	**F2 (Hz)**	**Std. Dev.**
avocado	838	89	1278	117
bratwurst	848	85	1518	225
Colorado	857	96	1393	125
drama	827	86	1296	247
façade	835	119	1275	140
Iran	770	124	1752	336
Iraq	908	97	1544	202
llama	837	115	1247	156
nachos	849	121	1414	138
pajamas	814	146	1488	235
panache	913	156	1598	151
pasta	915	122	1534	121
taco	810	103	1356	187

	Niagara, NY			
Word	**F1 (Hz)**	**Std. Dev.**	**F2 (Hz)**	**Std. Dev.**
avocado	862	95	1428	103
bratwurst	861	118	1571	188

Colorado	863	64	1389	108
drama	873	98	1373	127
façade	923	83	1412	92
Iran	696	155	1855	317
Iraq	848	117	1614	209
llama	930	76	1357	82
nachos	948	93	1552	119
pajamas	740	179	1728	311
panache	913	116	1512	109
pasta	944	85	1427	101
taco	950	88	1487	105

E. ANOVA and post-hoc Tukey test results

Questionnaire:

*Note: values rounded to three decimal places.

Overall aggregate:

ANOVA (one-way): Canadian Usage Aggregate					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	19487	9743	940.1	$<2e^{-16}$
Residuals	800	8291	10		
Tukey HSD (one-way): Canadian Usage Aggregate					
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
GTA-NON	0.835	-0.034	1.705	0.063	
NNY-NON	-10.100	-10.677	-9.523	0.000	
NNY-GTA	-10.935	-11.840	-10.031	0.000	
ANOVA (one-way): American Usage Aggregate					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	18884	9442	752.3	$<2e^{-16}$
Residuals	800	10040	13		
Tukey HSD (one-way): American Usage Aggregate					
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
GTA-NON	-0.712	-1.67	0.245	0.188	
NNY-NON	9.966	9.331	10.601	0.000	
NNY-GTA	10.678	9.683	11.674	0.000	

Pronunciation:

ANOVA (one-way): Canadian Usage Pronunciation					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	2497	1248.7	281.9	$<2e^{-16}$
Residuals	800	3544	4.4		

Tukey HSD (one-way): Canadian Usage Pronunciation				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	0.58	0.011	1.148	0.044
NNY-NON	-3.552	-3.929	-3.174	0.000
NNY-GTA	-4.132	-4.723	-3.54	0.000

ANOVA (one-way): American Usage Pronunciation					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	2221	1110.6	194	$<2e^{-16}$
Residuals	800	4581	5.7		

Tukey HSD (one-way): American Usage Pronunciation				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	-0.663	-1.309	-0.016	0.043
NNY-NON	3.321	2.892	3.75	0.000
NNY-GTA	3.983	3.311	4.656	0.000

Vocabulary:

ANOVA (one-way): Canadian Usage Vocabulary					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	8944	4472	1437	$<2e^{-16}$
Residuals	800	2490	3		

Tukey HSD (one-way): Canadian Usage Vocabulary				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	0.181	-0.296	0.658	0.646
NNY-NON	-6.92	-7.236	-6.604	0.000
NNY-GTA	-7.101	-7.597	-6.605	0.000

ANOVA (one-way): American Usage Vocabulary					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	9023	4511	1180	$<2e^{-16}$
Residuals	800	3058	4		

Tukey HSD (one-way): American Usage Vocabulary				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	-0.036	-0.564	0.492	0.986
NNY-NON	6.977	6.627	7.328	0.000
NNY-GTA	7.013	6.464	7.563	0.000

Spelling:

ANOVA (one-way): Canadian Usage Spelling					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	363.0	181.52	296.3	$<2e^{-16}$
Residuals	800	490.1	0.61		

Tukey HSD (one-way): Canadian Usage Spelling				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	-0.104	-0.316	0.107	0.477
NNY-NON	-1.418	-1.558	-1.278	0.000
NNY-GTA	-1.314	-1.534	-1.094	0.000

ANOVA (one-way): American Usage Spelling					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	393.3	196.7	281.6	$<2e^{-16}$
Residuals	800	558.7	0.7		

Tukey HSD (one-way): American Usage Spelling				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	0.034	-0.192	0.26	0.934
NNY-NON	1.464	1.314	1.614	0.000
NNY-GTA	1.430	1.195	1.665	0.000

Grammar:

ANOVA (one-way): Canadian Usage Grammar					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	161.6	80.81	118.7	$<2e^{-16}$
Residuals	800	544.8	0.68		

Tukey HSD (one-way): Canadian Usage Grammar				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	0.032	-0.191	0.255	0.941
NNY-NON	-0.929	-1.077	-0.781	0.000
NNY-GTA	-0.960	-1.192	-0.728	0.000

ANOVA (one-way): American Usage Grammar					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	143.1	71.55	81.53	$<2e^{-16}$
Residuals	800	702.1	0.88		

Tukey HSD (one-way): American Usage Grammar				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
GTA-NON	-0.098	-0.351	0.156	0.638
NNY-NON	0.86	0.692	1.028	0.000
NNY-GTA	0.957	0.694	1.221	0.000

Acoustic Study:

*Note: order of Tukey test pairings differs from questionnaire results. Values rounded to three decimal places.

/æ/-tensing:

ANOVA (one-way): /æ/ F1					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	3900440	1950220	214.5	$<2e^{-16}$
Residuals	450	4090995	9091		

Tukey HSD (one-way): /æ/ F1				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	-259.289	-293.812	-224.766	0.000
NON-GTA	-34.764	-63.007	-6.522	0.011
NON-NNY	224.524	197.170	251.879	0.000
ANOVA (one-way): /æ/ F2				

	df	*SS*	*MS*	*F*	*p*-value
Region	2	6539993	3269997	173.6	$<2e^{-16}$
Residuals	450	8474201	18832		

Tukey HSD (one-way): /æ/ F2				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	300.188	250.501	349.876	0.000
NON-GTA	-4.480	-45.127	36.167	0.964
NON-NNY	-304.668	-344.038	-265.299	0.000

Low back merger:

ANOVA (one-way): NON /ɑ/-/ɔ/ F1					
	df	*SS*	*MS*	*F*	*p*-value
Vowel	1	38832	38832	5.789	0.016
Residuals	810	5433384	6708		
ANOVA (one-way): NON /ɑ/-/ɔ/ F2					
	df	*SS*	*MS*	*F*	*p*-value
Vowel	1	241	32241	0.02	0.887
Residuals	810	9717080	11996		
ANOVA (one-way): NNY /ɑ/-/ɔ/ F1					
	df	*SS*	*MS*	*F*	*p*-value
Vowel	1	679645	679645	72.5	$1.67e^{-15}$
Residuals	246	2305996	9374		
ANOVA (one-way): NNY /ɑ/-/ɔ/ F2					
	df	*SS*	*MS*	*F*	*p*-value
Vowel	1	2668471	2668471	173.5	$<2e^{-16}$
Residuals	246	3782815	15377		
ANOVA (one-way): GTA /ɑ/-/ɔ/ F1					
	df	*SS*	*MS*	*F*	*p*-value
Vowel	1	3864	3864	0.511	0.475
Residuals	233	1760855	7557		
ANOVA (one-way): GTA /ɑ/-/ɔ/ F2					
	df	*SS*	*MS*	*F*	*p*-value
Vowel	1	4706	4706	0.378	0.54
Residuals	233	2904554	12466		

Low-Back-Merger Shift:

ANOVA (one-way): LBMS Index					
	df	*SS*	*MS*	*F*	*p*-value
Region	2	861640	430820	6.889	0.001
Residuals	195	12194236	62535		

Tukey HSD (one-way): LBMS Index				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	-186.226	-322.729	-49.723	0.004
NON-GTA	-29.314	-141.229	82.601	0.810
NON-NNY	156.912	48.378	265.446	0.002

Allophones of /æ/ (formant values):

ANOVA (one-way): NON /æ/ allophones F1					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	2	9426899	4713450	484.5	$<2e^{-16}$
Residuals	687	6683626	9729		

Tukey HSD (one-way): NON /æ/ allophones F1				
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
/æG/-/æ/	-199.486	-219.731	-179.241	0.000
/æN/-/æ/	-276.001	-298.766	-253.235	0.000
/æN/-/æG/	-76.514	-99.969	-53.059	0.000

ANOVA (one-way): NON /æ/ allophones F2					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	2	43173110	21586555	676	$<2e^{-16}$
Residuals	687	21939271	31935		

Tukey HSD (one-way): NON /æ/ allophones F2				
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
/æG/-/æ/	424.037	387.357	460.716	0.000
/æN/-/æ/	592.579	551.333	633.825	0.000
/æN/-/æG/	168.543	126.047	211.038	0.000

ANOVA (one-way): NNY /æ/ allophones F1					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	2	604574	302287	26.81	$4.26e^{-11}$
Residuals	210	2367592	11274		

Tukey HSD (one-way): NNY /æ/ allophones F1				
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
/æG/-/æ/	-5.654	-44.902	33.593	0.938
/æN/-/æ/	-129.072	-173.747	-84.398	0.000
/æN/-/æG/	-123.418	-169.337	-77.499	0.000

ANOVA (one-way): NNY /æ/ allophones F2					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	2	1999711	999856	34.02	$1.59e^{-13}$
Residuals	210	6171598	29389		

Tukey HSD (one-way): NNY /æ/ allophones F2				
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
/æG/-/æ/	45.799	-17.567	109.165	0.205
/æN/-/æ/	246.472	174.343	318.6	0.000
/æN/-/æG/	200.673	126.536	274.81	0.000

ANOVA (one-way): GTA /æ/ allophones F1					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	2	2855784	1427892	1118.8	$<2e^{-16}$
Residuals	195	2343870	12020		

Tukey HSD (one-way): GTA /æ/ allophones F1				
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
/æG/-/æ/	-183.889	-226.144	-141.634	0.000
/æN/-/æ/	-295.601	-343.08	-248.123	0.000

/æN/-/æG/	-111.712	-160.541	-62.883	0.000

ANOVA (one-way): GTA /æ/ allophones F2					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	2	9670156	4835078	159.5	$<2e^{-16}$
Residuals	195	5910596	30311		

Tukey HSD (one-way): GTA /æ/ allophones F2				
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
/æG/-/æ/	375.944	308.843	443.045	0.000
/æN/-/æ/	523.173	447.777	598.568	0.000
/æN/-/æG/	147.229	69.689	224.769	0.000

Allophones of /æ/ (Euclidean distance):

ANOVA (one-way): /æ/-/æG/ Euclidean Distance					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	1243302	621651	43.19	$1.54e^{-12}$
Residuals	63	906748	14393		

Tukey HSD (one-way): /æ/-/æG/ Euclidean Distance				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	-301.274	-416.654	-186.095	0.000
NON-GTA	52.003	-42.512	146.517	0.389
NON-NNY	353.377	261.718	445.036	0.000

ANOVA (one-way): /æ/-/æN/ Euclidean Distance					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	1341717	670858	19.15	$3.19e^{-7}$
Residuals	63	2207417	25038		

Tukey HSD (one-way): /æ/-/æN/ Euclidean Distance				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	-330.384	-510.251	-150.518	0.000
NON-GTA	24.443	-113.025	181.911	0.841
NON-NNY	364.827	221.814	507.840	0.000

ANOVA (one-way): /æG/-/æN/ Euclidean Distance					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	26607	13304	0.685	0.508
Residuals	63	1223791	19425		

Index of Phonetic Innovation:

ANOVA (one-way): IPI Values					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	4694061	2347030	60.01	$2.57e^{-15}$
Residuals	63	2464009	39111		

Tukey HSD (one-way): IPI Values				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	721.252	531.219	911.285	0.000
NON-GTA	69.705	-86.099	225.508	0.534
NON-NNY	-651.547	-802.644	-500.251	0.000

/u:/-fronting:

ANOVA (one-way): /u:/-fronting F2					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	5288066	2644033	26.74	$2.89e^{-11}$
Residuals	254	25113913	98874		
Tukey HSD (one-way): /u:/-fronting F2					
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
NNY-GTA	-402.630	-554.954	-250.306	0.000	
NON-GTA	-56.378	-181.295	68.538	0.537	
NON-NNY	346.252	226.323	466.181	0.000	

Canadian Raising (formant values):

ANOVA (one-way): NON Canadian Raising F1					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	3	4160753	1386918	166.8	$<2e^{-16}$
Residuals	1055	8771397	8314		
Tukey HSD (one-way): NON Canadian Raising F1					
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
/aiT/-/ai/	-118.855	-137.852	-99.857	0.000	
/auT/-/au/	-135.509	-157.957	-113.061	0.000	
/au/-/ai/	15.968	-5.946	37.881	0.239	
/auT/-/aiT/	-0.686	-20.298	18.925	0.9997	
ANOVA (one-way): NON Canadian Raising F2					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	3	9040432	3013477	141.3	$<2e^{-16}$
Residuals	1055	22505994	21333		
Tukey HSD (one-way): NON Canadian Raising F2					
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
/aiT/-/ai/	211.14	180.709	241.570	0.000	
/auT/-/au/	49.302	13.344	85.26	0.002	
/au/-/ai/	170.533	135.431	205.634	0.000	
/auT/-/aiT/	8.695	-22.72	40.109	0.892	
ANOVA (one-way): NNY Canadian Raising F1					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	3	1323477	441159	40.54	$<2e^{-16}$
Residuals	318	3460369	10882		
Tukey HSD (one-way): NNY Canadian Raising F1					
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
/aiT/-/ai/	-155.093	-194.900	-115.286	0.000	
/auT/-/au/	-51.638	-97.995	-5.281	0.022	
/au/-/ai/	-12.185	-57.728	33.359	0.901	
/auT/-/aiT/	91.271	50.535	132.006	0.000	
ANOVA (one-way): NNY Canadian Raising F2					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	3	2812202	937401	47.37	$<2e^{-16}$

Residuals	300	5936510	19788		
Tukey HSD (one-way): NNY Canadian Raising F2					
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
/aiT/-/ai/	203.767	148.486	259.048	0.000	
/auT/-/au/	86.876	22.639	151.113	0.003	
/au/-/ai/	163.429	100.437	226.422	0.000	
/auT/-/aiT/	46.538	-10.157	103.233	0.149	
ANOVA (one-way): GTA Canadian Raising F1					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	3	1030964	343655	34.26	$<2e^{-16}$
Residuals	300	3009091	10030		
Tukey HSD (one-way): GTA Canadian Raising F1					
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
/aiT/-/ai/	-108.498	-147.856	-69.140	0.000	
/auT/-/au/	-127.472	-173.206	-81.738	0.000	
/au/-/ai/	18.288	-26.56	63.136	0.718	
/auT/-/aiT/	-0.686	-41.050	39.678	0.9999695	
ANOVA (one-way): GTA Canadian Raising F2					
	df	***SS***	***MS***	***F***	***p*-value**
Vowel	3	2812202	937401	47.37	$<2e^{-16}$
Residuals	300	5936510	19788		
Tukey HSD (one-way): GTA Canadian Raising F2					
Vowel Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
/aiT/-/ai/	203.767	148.486	259.048	0.000	
/auT/-/au/	86.876	22.639	151.113	0.003	
/au/-/ai/	163.429	100.437	226.422	0.000	
/auT/-/aiT/	46.538	-10.157	103.233	0.149	

Canadian Raising (Euclidean distance):

ANOVA (one-way): /au/-/auT/ Euclidean Distance					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	99769	49885	7.718	0.001
Residuals	63	407182	6463		
Tukey HSD (one-way): /au/-/auT/ Euclidean Distance					
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
NNY-GTA	-105.341	-182.592	-28.091	0.005	
NON-GTA	-10.443	-73.779	52.892	0.917	
NON-NNY	94.898	33.475	156.320	0.001	
ANOVA (one-way): /ai/-/aiT/ Euclidean Distance					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	62311	31155	4.279	0.018
Residuals	63	458733	7281		
Tukey HSD (one-way): /ai/-/aiT/ Euclidean Distance					
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**	
NNY-GTA	-69.003	-150.998	12.992	0.116	

NON-GTA	9.934	-57.292	77.16	0.933
NON-NNY	78.937	13.743	144.132	0.014

Foreign (a) (overall):

ANOVA (one-way): Foreign (a) F1					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	468134	234067	18.93	$9.11e^{-9}$
Residuals	841	10400314	12367		

Tukey HSD (one-way): Foreign (a) F1				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	24.896	-4.323	54.115	0.113
NON-GTA	-32.912	-56.904	-8.919	0.004
NON-NNY	-57.808	-81.009	-34.606	0.000

ANOVA (one-way): Foreign (a) F2					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	1798889	899445	15.6	$2.23e^{-7}$
Residuals	841	48493126	57661		

Tukey HSD (one-way): Foreign (a) F2				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	76.905	13.811	139.999	0.012
NON-GTA	-41.581	-93.39	10.227	0.144
NON-NNY	-118.486	-168.586	-68.387	0.000

Foreign (a) (pasta):

ANOVA (one-way): Foreign (a) - pasta F1					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	130452	65226	9.381	0.0003
Residuals	61	424130	6953		

Tukey HSD (one-way): Foreign (a) - pasta F1				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	29.264	-52.797	111.324	0.669
NON-GTA	-75.563	-143.759	-7.368	0.026
NON-NNY	-104.827	-168.776	-40.878	0.0006

ANOVA (one-way): Foreign (a) - pasta F2					
	df	***SS***	***MS***	***F***	***p*-value**
Region	2	491722	245861	9.77	0.0002
Residuals	61	1535004	25164		

Tukey HSD (one-way): Foreign (a) - pasta F2				
Region Pairing	**mean difference**	**lower bound**	**upper bound**	***p*-value**
NNY-GTA	-106.412	-262.526	49.701	0.238
NON-GTA	-225.817	-355.553	-96.081	0.0003
NON-NNY	-119.405	-241.062	2.253	0.055