Grammatical Variation and Change in Industrial Cape Breton

by

Matthew Hunt Gardner

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto

© Copyright by Matthew Hunt Gardner (2017)
Abstract

Grammatical Variation and Change in Industrial Cape Breton

Matthew Hunt Gardner
Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto
2017

The following dissertation explores four changes in progress in a large mixed corpus of speech from Cape Breton Island. Its goal is to establish evidence for a generational genetic relationship between Cape Breton English and Scottish/Irish English, if any. Findings indicate that Cape Breton English is instead more similar to Inland Canadian English with respect to both quantity and kind for these changes. This paper also explores the processes of transmission and diffusion on Canadian English’s eastern edge. It contributes to the growing body of literature on change in the English stative possessive, deontic modality, future temporal reference, and quotative systems.
for Joan
Acknowledgements

I would like to thank my advisors Sali A. Tagliamonte, Jack Chambers, and Aaron Dinkin for their support, feedback, encouragement, and especially patience during this project. I would also like to thank Jack Chambers, Gerard Van Herk, and Paul De Decker for their guidance during data collection. I would not have been able to write this dissertation without extended conversations with my peers, including Derek Denis, Marisa Brook, Bridget Jankowski, Rebecca Roeder, and Becky Childs. I would also like to thank the full Language Variation and Change Research Group at the University of Toronto and the Memorial University Linguistics Department for feedback on earlier versions of this research.

This project would contain much less data if it were not for Jane Arnold, archivist at the Beaton Institute at Cape Breton University. Her willingness to track down a computer drive capable of reading 5.25 floppy disks from 1990 meant I was able to include the Steelworker corpus in my research. I must also thank Jessie Fraser, Bridget Henley, and especially Kathleen Power for their assistance processing the Storyteller and Post-Industrial corpora.

Susan and Karl Misik, George, Bill, Jean, Jenna and Chris Gardner tried their best to feign interest in my accounts of the origins of Cape Breton English. They also were instrumental in helping me collect the Post-Industrial corpus. Jenna Gardner was especially helpful — drinking endless cups of tea in Cape Breton kitchens with me as I interviewed people. Thanks sis.

Edison, you’re my best friend and you more than anyone else kept me sane and laughing over the last seven years. Thank you.

Finally, this research was supported by the Social Sciences and Humanities Research Council of Canada and the following funding sources: the Thomas and Beverly Simpson Ontario Graduate Scholarship; the Ontario Graduate Scholarship program; the University of Toronto Linguistics Department; the University of Toronto School of Graduate Studies; and Karl and Susan Misik.
## Contents

Acknowledgements iv  
Table of Contents v  
List of Tables vii  
List of Figures ix  
Table of Photographs xi  
List of Maps xii

### I  Introduction

1 Introduction 2  
1.1 Outline of the current study 4  
1.2 The variables 5  
1.3 The community 6  
1.4 Relationship to other work 7

2 Sociolinguistic Issues 9  
2.1 Research framework 9  
2.2 Changes through time and space 9  
2.3 Peripherality and diachrony 12  
2.4 Layering versus functional distinctions 13  
2.5 Age, Sex and Social Status 14  
2.6 Summary 15

3 Cape Breton English 16  
3.1 The history of Cape Breton and its English 17  
3.1.1 Pre-Industrial Cape Breton (late 1700s–late 1800s) 17  
3.1.2 Cape Breton English in the 20th-Century 22  
3.2 Previous Accounts of Cape Breton English 27

### II  Methods

4 The Data 38  
4.1 The McKinnon Texts 39  
4.2 Storyteller corpus 41  
4.3 Steelworker corpus 42  
4.4 Corpus of Post-Industrial Cape Breton English 43
5 Age, Sex, Social Status, etc. 44
5.1 Age 44
5.2 Sex 46
5.3 Social Status 48
5.4 Race, ethnicity, religion, etc. 49

III Case Studies 50

6 Quotatives 51
6.1 Introduction 51
6.2 Synchronic perspective 52
6.3 Methods 53
6.4 Results 54
6.4.1 Trajectories of Change 54
6.4.1.1 All Quotatives 54
6.4.1.2 Go 55
6.4.1.3 Be like 57
6.4.2 Linguistic Conditioning 61
6.5 Discussion 63
6.6 Conclusions 70

7 Stative Possessives 72
7.1 Introduction 72
7.1.1 The Variable 72
7.2 Historical Perspective 74
7.2.1 The McKinnon Texts 75
7.3 Synchronic perspective 76
7.4 Methods 77
7.5 Results 78
7.5.1 Stative possessive verbs and do-support 78
7.5.2 Trajectories of Change 81
7.5.3 Linguistic Conditioning 83
7.5.3.1 Subject Type 83
7.5.3.2 Object Type 84
7.5.3.3 Subject Reference 84
7.5.3.4 Sex 85
7.5.3.5 Logistic Regression Analysis 85
7.6 Discussion 90
7.6.1 have/have got/got 90
7.6.2 Have got/got 91
7.7 Conclusion 96

8 Deontic Modality 97
8.1 Introduction 97
8.2 The Variable 97
8.3 Historical Perspective 98
8.3.1 The McKinnon Texts 99
8.4 Synchronic Perspective 100
8.5 Methods 103
8.6 Results 104
8.6.1 Deontic Modality and Discourse 104
8.6.2 Deontic and Epistemic Modality 106
8.6.3 Deontic modality verbs and do-support 107
List of Tables

2.1 Accounting for similar changes across varieties ............................................. 11
3.1 Linguistic stratification of Cape Breton, early 1800s ................................. 20
3.2 Principal Vowels of Cape Breton English .................................................... 32
3.3 Principal Vowels of Standard Scottish English ............................................. 32
3.4 Gaelic-origin borrowings/calques in Cape Breton English .......................... 33
4.1 Tokens extracted from the McKinnon Texts ................................................ 40
4.2 Tokens extracted from the Storyteller corpus ............................................ 42
4.3 Tokens extracted from the Steelworker corpus .......................................... 43
4.4 Tokens extracted from the Post Industrial corpus ..................................... 43
6.1 Quotation in Cape Breton English .............................................................. 55
6.2 Quotation in Cape Breton English .............................................................. 57
6.3 Analysis of deviance .................................................................................. 62
6.4 Mixed effects logistic regression ............................................................... 64
6.5 Distribution of it’s like across corpora ....................................................... 66
7.1 Distribution of stative possessive verbs in Cape Breton English ................. 78
7.2 Distribution of stative possessive verbs in Cape Breton English ............... 79
7.3 Distribution of stative possessive verbs in Cape Breton English ............... 80
7.4 Distribution of stative possessive verbs in Cape Breton English ............... 81
7.5 Linguistic and social constraints governing the use of have for stative possession across varieties 85
7.6 Analysis of deviance .................................................................................. 86
7.7 Analysis of deviance .................................................................................. 88
7.8 Mixed effects logistic regression ............................................................... 89
7.9 Stative possessive got vs. have got in Cape Breton and Toronto English .... 93
8.1 Distribution of deontic modality verbs in Cape Breton English ................. 105
8.2 Distribution of deontic modality verbs in Cape Breton English ............... 105
8.3 Distribution of epistemic modality verbs in Cape Breton English ............. 106
8.4 Distribution of deontic modality verbs in Cape Breton English ............... 107
8.5 Distribution of deontic modality verbs in Cape Breton English ............... 108
8.6 Distribution of deontic modality verbs in Cape Breton English ............... 109
8.7 Analysis of deviance .................................................................................. 116
8.8 Mixed effects logistic regression ............................................................... 118
9.1 Distribution of expressions of future temporal reference in early Cape Breton English ................................................................. 126
9.2 Distribution of expressions of future temporal reference in Cape Breton English ................................................................. 133
9.3 Distribution of expression of future temporal reference in Cape Breton English .................................................................................. 141
9.4 Mixed effects logistic regression of the contribution of factors to the probability of be going to for expressing future temporal reference across corpora and age groups in Cape Breton English. 150
A.1 Mixed effects logistic regression ............................................................... 201
## List of Figures

1.1 Population of Cape Breton County/CBRM 1851 to 2011 ........................................... 7

5.1 Speakers by sex, corpus and year of birth ................................................................. 45
5.2 Tokens by corpus and year of birth .............................................................................. 45
5.3 Speakers by sex and corpus ......................................................................................... 47
5.4 Speakers by sex, corpus and year of birth ................................................................. 47

6.1 Distribution of quotative verbs in Cape Breton English ............................................ 54
6.2 Distribution of *be like* and *go* across corpora of Cape Breton English ................. 56
6.3 Distribution of *be like* by content of the quotation in Cape Breton English .......... 58
6.4 Distribution of *be like* by grammatical person of the subject in Cape Breton English 60
6.5 Distribution of *be like* by tense/temporal reference in Cape Breton English .......... 61
6.6 Distribution of *be like* by sex in Cape Breton English ............................................. 65
6.7 Distribution of *be like* by subject person and tense/temporal reference in Cape Breton English 67
6.8 Distribution of *be like*, including *it’s like* and *I’m like* in Cape Breton English .... 69

7.1 Distribution of stative possessive *have* in Canadian and British English .................... 76
7.2 Distribution of stative possessive verbs in conservative Scottish and Irish English .......... 77
7.3 Distribution of stative possessive verbs in Cape Breton English ................................ 82
7.4 Distribution of stative possessive verbs in Cape Breton, Toronto and Buckie ............. 82
7.5 Distribution of stative possessive verbs in Cape Breton English .............................. 84
7.6 Conditional inference recursive partitioning tree ......................................................... 87
7.7 Conditional inference recursive partitioning tree ......................................................... 92

8.1 Distribution of deontic modality verbs in conservative Scottish and Irish English .......... 102
8.2 Distribution of deontic modality verbs in African–Nova Scotian English ....................... 102
8.3 Distribution of deontic modality verbs in Cape Breton English ................................ 109
8.4 Secondary deontic modality verbs across corpora of Cape Breton English ............... 110
8.5 Distribution of deontic modality verbs in Cape Breton, Toronto and Buckie ............. 110
8.6 Distribution of deontic modality verbs in Cape Breton English ................................ 111
8.7 Distribution of deontic modality verbs in Cape Breton English ................................ 112
8.8 Distribution of deontic modality verbs in Cape Breton English .............................. 113
8.9 Distribution of deontic modality verbs in Cape Breton English .............................. 114
8.10 Conditional inference recursive partitioning tree ...................................................... 115
8.11 Conditional inference recursive partitioning tree ...................................................... 116
8.12 Distribution of deontic modality verbs in Cape Breton English .............................. 119
8.13 Distribution of stative possessive *have* and deontic modality *have to* in Cape Breton English 120

9.1 Distribution of *be going to* in conservative varieties of Scottish and Irish English English 131
9.2 Distribution of expression of future temporal reference in Cape Breton English .......... 132
9.3 Distribution of *be going to* across varieties of English ........................................... 133
9.4 Distribution of *be going to* by sentence type in Cape Breton English ....................... 136
9.5 Distribution of *be going to* by proximity in Cape Breton English ........................... 138
9.6 Conditional inference recursive partitioning tree ...................................................... 140
# List of Photographs

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cape Breton T-shirt</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Cape Breton T-shirt</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>Newfoundland T-shirt</td>
<td>36</td>
</tr>
<tr>
<td>4</td>
<td>Better Health Centre Advertisement</td>
<td>73</td>
</tr>
<tr>
<td>5</td>
<td>Woolco Advertisement</td>
<td>73</td>
</tr>
<tr>
<td>6</td>
<td>Who ya got?</td>
<td>96</td>
</tr>
<tr>
<td>7</td>
<td><em>Must</em> for expressing deontic modality in signage</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td><em>Must</em> for expressing deontic modality in signage</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td><em>Must</em> for expressing deontic modality in signage</td>
<td>101</td>
</tr>
</tbody>
</table>
## List of Maps

<table>
<thead>
<tr>
<th></th>
<th>Map Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cape Breton Island</td>
<td>xiii</td>
</tr>
<tr>
<td>2</td>
<td>Industrial Cape Breton</td>
<td>xiv</td>
</tr>
<tr>
<td>3</td>
<td>Cape Breton Island, selected communities</td>
<td>xv</td>
</tr>
<tr>
<td>4</td>
<td>Nova Scotia</td>
<td>xvi</td>
</tr>
<tr>
<td>5</td>
<td>Southern Ontario</td>
<td>xvii</td>
</tr>
<tr>
<td>6</td>
<td>The British Isles</td>
<td>xviii</td>
</tr>
</tbody>
</table>
Map 1: Cape Breton Island. Adapted from "Cape Breton Island" by Klaus Mikmaq, Germany, January 3 2007. CC image courtesy of Klaus Mikmaq, via Wikimedia Commons.
Map 2: Industrial Cape Breton (large red circle), and Port Hawkesbury Industrial Area (small red circle) (Cape Breton Development Corporation 1968, inside front cover).
Map 3: Cape Breton Island, selected communities.
Map 5: Southern Ontario, including Toronto and the three communities studied by Tagliamonte and Denis (2014): Burnt River, Lakefield, and Belleville.
Part I

Introduction
Chapter 1

Introduction

The Cape Breton speech community has been described as a strongly recessive relic speech enclave showing signs of convergence with Inland Canadian English (Boberg 2010, 27; Dollinger 2008, 16; Keifte and Kay-Raining Bird 2010, 62). The divergent features of Cape Breton English are often attributed — by linguists and others — to the island’s distinctive input population (specifically Scottish Gaels) and history of geographic, social, and economic isolation. For example:

The distinctive quality of Maritime speech phonetically comes down largely to its treatment of [ɑ] before medial and final [r] and occasionally elsewhere, which is somewhat more fronted than in General American — somewhat similar to the New England treatment of this sound. Otherwise, Maritime speech is similar to Canadian English (or General American) in pronunciation, intonation and vocabulary (expect for a few localism which can probably be traced to New England). In Scottish-derived areas of Cape Breton Island and Pictou County, however, a number of Gaelic words are to be found in ordinary English speech, (M. Bloomfield 1948, 8).

Although, it has not been studied extensively, the speech of Cape Breton seems to bear many similarities with the nearby island of Newfoundland and Westerners can perhaps be excused for their inability to differentiate the two accents. These similarities — e.g. fronting of the low back vowel — owe much to the geographic proximity of the two islands, the fact that approximately one-quarter of the Cape Breton population descends from Irish immigrants — many of whom arrived via Newfoundland — and the Celtic influences in each region (i.e., either Scottish or Irish), (Keifte and Kay-Raining Bird 2010, 68).

The British accents and dialects of the 19th century immigrants made a direct and indisputable impression on Canadian speech. First, in relatively isolated regions where the immigrants became the founding population, their speech formed the basis of the local accent. To this day, one can discern the Scots roots of rural speech in Cape Breton, Pictou and Antigonish counties in Nova Scotia, the Ottawa Valley, Peterborough county, the West Lorne district on the north shore of Lake Erie, and other places, (Chambers 2004a, 228).

1. Inland Canadian English, used here and throughout, describes the dialect of English spoken in Canada from Ontario westward. The label "Inland Canadian English" was coined by Labov, Ash, and Boberg (2006, Section 15.6) by analogy with the American Inland North. In its original conception Inland Canada is more restricted than Canadian English, as it does not include speakers from Vancouver or Angophone Québec. Here I include these two regions within what I call Inland Canada because this term with this definition is a useful heuristic for discussing dialect differences between Atlantic Canada and the rest of the country. Inland Canadian English is relatively synonymous with Mainstream Canadian English and Standard Canadian English, though these terms additionally contrast with other more specific regional or ethnic varieties (e.g., Ottawa Valley English, Montréal Jewish English, Toronto Jamaican English, Indigenous Canadian English, etc.) which are, like Canadian Maritime English and Newfoundland English, viewed as divergent from the mainstream/standard.
Cape Breton Island English, Nova Scotia, shows considerable influence from Scottish Gaelic... My own fieldwork on the English of Cape Breton, Nova Scotia, reveals a variety that is more or less identical with the English of the Scottish Highlands, (Trudgill 2004, 5, 7).

The eastern provinces [of Canada] are well known for their dialect pockets and the range of language contacts: from the influence of Irish English in Newfoundland, to Scottish Gaelic in Cape Breton, French influence in New Brunswick, and German and New England influence in Nova Scotia, (Gold 2010, 28).

A quick search of Google Books shows that when characters or real-life individuals are described as having a Cape Breton accent, that accent is nearly always qualified as being divergent (e.g., “strong”, “thick”), humorous, or as being “lilting” or a “brogue” — traditional descriptors of Gaelic, Irish, and Scottish/Irish English. Several examples are listed below, but the list could be greatly extended.

[Of singer-songwriter, Rita MacNeil] Her Cape Breton accent gives a soft lilt and musicality to her voice. While she is friendly, and answers every question that is asked of her, there is an awkwardness that reveals her shyness when forced to share her thoughts through speech, (Rhindress 2017, 100).

[Of folk singer John Allan Cameron] Cameron’s sound, of course, was important — the gently lilting Cape Breton accent, the 12-string guitar — but even more important was the personality that shone on stage, something which no audio recording can ever quite capture, (MacDonald 2012, 5).

“Now how do you suppose the victim knew your dog?” asked Mansour with the hint of a lilting Cape Breton accent, (Kid 2012, 164).

“Well don’t that give your toque a spin,” William said. “I got girls falling on me like summer ticks.” When I heard that flat drawn out Cape Breton accent I nearly laughed so hard I almost peed myself, (Vernon 2009, 247).

Billy thought he was pretty good at teasing people but he was an amateur compared to Bill McDonald, the Nova Scotian. McDonald was forever pulling pranks on everyone and constantly teasing Ma. He had a thick Cape Breton accent that amused Billy. Words like “sure” came out as “shore,” and with his huge jug ears and lantern jaw, he kept everyone in stitches, (McLean 2011, 100).

“I suppose you’re going to blame the fairies,” my aunt said. The lingering trace of a Cape Breton accent always came out when she was angry. It exploded into an unmistakable brogue just now, (Lee 2017).

He was the only man in his troop from Nova Scotia and was constantly reminded about his Cape Breton accent with laughable and sometimes irritating imitations, (Smith 2012, 17).

Modern travel guides prepare prospective travellers to Cape Breton to hear the island’s distinct, Gaelic-origin variety.

In Chéticamp, an Acadian enclave for more than 200 years, Francophone culture and traditions are still very much alive. That’s why the Gaelic-inflected lilt in locals’ voices is replaced by a distinct French accent, (Fodor’s Travel 2017).
Aye, the Scottish influence is unmistakable. Upwards of fifty thousand Highlanders found their way to Cape Breton in the early nineteenth century, bringing their Gaelic language with them. Centuries later, discovering Scotland in Canada charms the kilt off most visitors, unaccustomed to the distinct Cape Breton accent they encounter in small communities dotting the island, (Esrock 2013, 61).

This classification of Cape Breton speech as a relic Scottish variety is, however, based on anecdotal evidence only, and has likely been left unchallenged because both being different from the rest of Canada and being Scottish are integral components of how the community self-identifies. This dissertation aims to test this claim using cutting edge statistical techniques. It will assess the structural similarities in variable grammatical patterns in a multimillion word corpus of Cape Breton English alongside published accounts of English spoken in Ontario and in Northern England, Northern Ireland and Scotland.

Cape Breton English sits on the extreme eastern periphery of Canadian English’s geographic range. For this reason the community also offers a prime opportunity to explore how linguistic changes diffuse eastward from large cities like Toronto and Montréal. Furthermore, the relatively large number of Scottish Gaelic- and Scottish/Irish English-speaking early settlers in the community makes Cape Breton English a good test-site for studying the Scots/Irish influence, if any, on Canadian English — a heretofore understudied topic according to Dollinger (2012, 1862) and Clarke (1997, 208).

While patterns of convergence with mainstream Canadian English have been observed quantitatively for both phonetic and phonological features in Cape Breton English (for example, Gardner 2010b; 2013a; Roeder and Gardner 2013; Keifte and Kay-Raining Bird 2010; and also Rowe 1968), no quantitative analysis of Cape Breton English above the level of phonology exists. This dissertation will serve as a preliminary document of grammatical patterning of this variety. This dissertation further presents a significant contribution to the scholarship of several major changes occurring in varieties of English, including changes to the stative possessive, deontic modality, and future temporal reference systems, as well as the rapid rise in use of quotative be like. The datasets for stative possession, deontic modality, and future temporal reference are the largest ever compiled within the variationist sociolinguistic literature. The quotative dataset is also the largest compiled for a single community other than Toronto. With such large datasets, this dissertation is able to offer both confirmation of the grammatical patterning of these changes found elsewhere, as well as new insights for how these patterns developed and spread.

1.1 Outline of the current study

This dissertation is organized as follows. In this first chapter of Part I the community of Cape Breton is introduced, as are the four sociolinguistic variables that will be used to assess linguistic transmission and diffusion in Cape Breton English. There is also a description of how this study relates to other work. Chapter 2 presents the sociolinguistic issues relevant to this study. Chapter 3 provides a history of the transplanting of English to the island of Cape Breton, with the aim of determining the precise roots of the variety. It also describes the sociolinguistic milieu on Cape Breton Island since the early 1800s, and how the island’s changing economic, demographic, and cultural history has shaped patterns of linguistic variation and change. Finally, Chapter 3 outlines past linguistic research on Cape Breton English and assesses the evidence presented by this past research for a connection between Cape Breton English and Scottish Gaelic/Scottish English.

In Part II, Chapter 4 outlines the four corpora of Cape Breton English that were used for modelling linguistic behaviour and Chapter 5 discusses how social factors like age, sex, and social status are operationalized.

Part III contains the four case studies used to assess grammatical transmission and diffusion in Cape Breton
Chapter 1. Introduction

English. The first of these case studies, Chapter 6, tracks the rise of quotative be like in Cape Breton English. This change is categorized as prototypical linguistic diffusion for this community, and serves as a point of comparison when categorizing the three following changes as either transmitted or diffused. Chapter 7 examines changes in Cape Breton for expressing stative possession; Chapter 8 examines the deontic modality system; and Chapter 9 examines expressions of future temporal reference.

Part IV is an overall review and discussion of the findings from the four case studies. Chapter 10 discusses trends that emerge when examining the results of the case studies together, while Chapter 11 summarizes the work presented in this paper and offers insight into the relevance of the major finding to the community of Cape Breton and the categorization of Cape Breton English.

Throughout this dissertation topics that warrant further or future study are highlighted. These include topics specific to Cape Breton English, the four variables under investigation, and language variation and change generally.

1.2 The variables

In order to explore linguistic transmission and diffusion in a peripheral context and possible Scottish influences, this dissertation will focus on four variables above the level of phonology: three long-term morphosyntactic changes occurring in the English language with differing grammatical patterning in Scotland and Inland Canada, and one very rapid discourse/pragmatic change that is diffusing worldwide.

The morpho-syntactic changes are:

1. The expression of present tense stative possession (have/have got/got), see Chapter 7.

   (1) It’s two stories, I think. It’s about 1,300 square feet, I think. And we [underlined] got about two-and-a-half acres of back yard — half of which my dad actually mows. And we [underlined] have a pool table in the basement. (Male born 1991, PI)  

2. The expression of deontic modality (must/have to/have got to/got to/need to), see Chapter 8.

   (2) Then we moved on to plastic. Plastic was what they were using at the last of it when I was there. You [underline]’ve got to tamp that in. You [underline] have to be sure. You [underline] have to work fast. (Male born 1921, SW)

3. The expression of future temporal reference (shall/will/be going to, simple present tense, and others.), see Chapter 9.

   (3) It’s designed to scare the public — well it’s to raise the money — but we’ll actually have like a— it’s going to be an old theatre in the school, that’s what we’re reopening, that’s the theme of it. So it’s going to be, like, a stage production of The Wizard of Oz and we’re going to take everybody in into the theatre. Dorothy is going to be singing Somewhere Over the Rainbow and then the music is going to start skipping and then she’ll turn around and she’ll be dead and then she might run off stage or something. (Male born 1991, PI)

The discourse/pragmatic change is:

4. Quotative verbs (say/think/go be like, and others.), see Chapter 6.

   (4) He looks like a kid who could be in junior high school. And I said to Elise — like, I didn’t even get the sentence out — I said “Is that…?” She [underline] goes, “Yeah, that’s him.” And I was like, “Does he…?” She [underline] goes, “…look like he’s a kid?” And I was like, “Yeah!” (Male born 1985, PI)

2. Here and throughout PI refers to the Post-Industrial corpus, see Chapter 4.
3. Here and throughout SW refers to the Steelworker corpus, see Chapter 4.
Beyond assessing the genetic link between Scottish Gaelic/Scottish English and Cape Breton English, each of these features was specifically chosen to test theories of linguistic change. For example, the use of *have got* or *got*, rather than *have*, for the expression of stative possession is particularly germane to the study of language change because it atypically reversed through history — at least in North America. The use of *be going to* rather than *will* for the expression of future temporal reference, on the other hand, is a prime variable for studying the relative isolation of a variety because the change has yet to penetrate or has only recently penetrated peripheral English-speaking communities. The different variants that express deontic modality entered the English language at specific documented points, so their distribution across a variety can shed light both on the evolution of the linguistic system and language change generally. Finally, the late-20th-century meteoric rise of the *be like* quotative has been tracked in real and apparent time in many English speaking communities. It is especially useful for testing theories of how linguistic variables spread from community to community and how lexical expressions take on grammatical meanings.

### 1.3 The community

The island of Cape Breton is located on the eastern extremity of the Gulf of St. Lawrence and is separated from mainland Nova Scotia by the narrow Strait of Canso, bridged by a 2 km causeway built in 1955, and from neighbouring Newfoundland by the 110 km-wide Cabot Strait (see Map 1, p. xiii). While currently part of the Canadian province of Nova Scotia, Cape Breton Island has for periods of its history been considered politically separate from the rest of the Nova Scotian peninsula.

The island’s current population, 131,292 (as of July 2015), represents about 14 percent of the total population of Nova Scotia (Milloy 2014). About three-quarters of the island’s population lives in the Cape Breton Regional Municipality (or CBRM), which comprises the entirety of Cape Breton County, itself the eastern quarter of the island. The CBRM was incorporated in 1995, amalgamating the Municipality of Cape Breton, the City of Sydney, and the Towns of Glace Bay, Sydney Mines, New Waterford, North Sydney, Dominion and Louisbourg. While considered one municipality nowadays, traditional boundary lines remain culturally important for community members who often identify their home communities by former community names. The official English motto of the CBRM is “a community of communities.”

The eastern part of the CBRM has been traditionally referred to as “Industrial Cape Breton” for the industrial operations in the area (see Map 2, p. xiv). The term is most often employed to distinguish the area from the rest of the island, which traditionally relied on fishing, farming and later tourism and pulp and paper manufacturing.

---

4. With the exceptions of the Eskasoni and Membertou First Nations communities
Between 1900 and 1962 Cape Breton County experienced very rapid industrialization with accompanying population growth and diversity. Since the community’s population peak in 1961, a 30-year period of de-industrialization spurred massive population decline in the community. The rise and fall of heavy industry in Cape Breton, and the demographic changes that were its consequence, have meant that speakers’ occupations, need for education, and income levels have all radically changed. Social ties have also changed, as first the community swelled with members, and then lost members through outmigration at increasingly larger rates. As the community changed, the way in which speakers could engage with and within the community changed too. Though these changes (especially the latter) are not unique to Cape Breton, their severity and swiftness make Cape Breton English a useful object of study that can help refine our knowledge of the relationship between socio-economic/socio-demographic/socio-cultural realities and language variation and change. This dissertation aims to explore whether the expected patterns of transmission and diffusion are present or are interrupted by massive out-migration. Do speakers rapidly converge to mainstream norms, or do they retreat to conservative ways of speaking?

1.4 Relationship to other work

My dissertation is part of the tradition of the University of Toronto Sociolinguistics/Language Variation and Change laboratory, which conducts systematic studies of language change, especially for Canadian English. I take as my model a collection of studies that have grown out of the laboratory’s research, including Tagliamonte and D’Arcy (2004), Tagliamonte (2006c), Tagliamonte and D’Arcy (2009), Gardner et al. (2013), Tagliamonte (2013a), Tagliamonte and Denis (2014), Tagliamonte, Durham, and Smith (2014), and Denis and Tagliamonte (2017). Collectively these studies present comparable analyses of the above four grammatical/discourse-pragmatic changes using data collected in both Southern Ontario and the northern United Kingdom, representing both the variety to which Cape Breton English is believed to be converging and those from which it is believed to be diverging. These studies also present testable claims about the way in which these four changes are progressing through time and from variety to variety.

For example, Tagliamonte (2006b) tracks the rise of innovative variants for each of these features in Toronto.
English and shows how the variable systems for these variants distinguish Toronto English from northern British English. Tagliamonte and Denis (2014) examine three of these variables (stative possession, deontic modality, and quotative verbs) in Toronto and in three communities outside of that city. The authors problematize the notion that these changes are being led by Toronto speakers and have (or are being) diffused to the outlying communities. Tagliamonte (2013a) examines the same three features (plus expressions of future temporal reference) among several peripheral speech communities in the British Isles and compares the distribution of the variants and the internal linguistic constraints governing their use among the communities and between the communities. The be like quotative has yet to penetrate the speech of the older speakers from the peripheral communities studied by Tagliamonte (2013a) so it is not included in that work; however, Tagliamonte and Hudson (1999) and Gardner et al. (2013) do examine the rise of this variant among younger speakers from York, U.K., one of the communities Tagliamonte (2013a) studies. Gardner et al. (2013) compares the rise of be like in York with its rise in Toronto, finding evidence of layering (to the point of saturation) in Toronto and clear evidence of diffusion in York. 5

This dissertation builds on the work of Tagliamonte and Denis (2014) by extending the potential range of diffusion to the easternmost edge of the Canadian English isogloss, and by testing whether the trends observed by the authors hold true over a greater distance and despite geographic and linguistic barriers. As Cape Breton is at the periphery of Canadian English, this dissertation will also build on the work of Tagliamonte (2013a) by observing whether peripherality begets conservatism in a Canadian context, and by determining whether the way in which these four variables permeate peripheral varieties occurs along a similar trajectory in Cape Breton as they do in the northern British Isles. Cape Breton English provides a new Canadian variety through which we can observe the rise of innovative variants (e.g., be like) and test claims of layering, specialization, grammaticalization and diffusion.

Finally, descriptions of Cape Breton English have consistently attributed its divergent characteristics to its conservation of specifically Scottish/Irish features. Comparing Cape Breton English data to Tagliamonte (2013a) and Tagliamonte, Durham, and Smith’s (2014) work on northern Irish, northern British, and Scottish varieties and Tagliamonte and Denis’s (2014) work on Ontario English shows quantitatively (for the first time) that the immediate roots of Cape Breton English cannot be the northern British Isles. Instead speakers of Cape Breton English align closely at the morphosyntactic and discourse-pragmatic level with Loyalist-rooted Inland Canadian English. This finding is quite provocative given that the historically-distinct varieties of the Maritimes are either ignored in discussions of Canadian English’s homogeneity (M. Bloomfield 1948, 63; Avis 1973, 50-51; Chambers 2006, 385) or excluded from being part of Canadian English at all. For example, Boberg (2010, 28) writes that “…the varieties of English found in these enclaves [like Cape Breton] can be thought of as distinct from Canadian English in the sense that they all represent the failure of initially non-Canadian varieties (mainly Scottish English, Irish English, and African American English) to assimilate fully to the Canadian English spoken in the larger regions around them.” If Cape Breton English is framed instead as peripheral, and thus an older form of Canadian English, rather than an enclave or island external to Canadian English, it instead offers insight into how Canadian English homogenized in the first place.

---

5. Though see Gardner et al. (2016) for an alternative interpretation of the same data
Chapter 2

Sociolinguistic Issues

The following section outlines how this research intersects with sociolinguistic theory. Presented is the theoretical underpinnings of the variationist sociolinguistic methods employed. Following this, several ways that this research adds insight into how, why, and where language changes is described.

2.1 Research framework

My point of departure is the linguistic variable and variable rules. At their most basic, the linguistic variable and variable rules are useful heuristic devices for describing when in a language there is more than one way of saying the same thing (Labov 1978, 1). More technically, the variable is an abstract construct which represents a group of options within the same grammatical system that have the same referential value in running discourse. (Sankoff 1988, 142-143; Tagliamonte 2012b, 4). The choice between options, or variants, of a variable is patterned, not random, and this patterned variation reflects the structured order of the grammar. Weinreich, Labov, and Herzog (1968, 100) call this “orderly heterogeneity.” The choice of variants also reflects the identity of the speaker (among other things). This is because language does not simply transmit information, but makes a statement about who the speaker is, with what group the speaker’s loyalties reside, how the speaker perceives his or her relationship to his or her hearers, and what sort of speech event the speaker considers him or herself to be engaged in (Tagliamonte 2006a, 7). Each variant of the four variables introduced in Section 1.2 may be engaged in language change but also occur more or less frequently with certain speakers, or certain kinds of speakers, and may be indexically linked to certain identities, contexts, or styles of speech. A variable rule is simply the probabilistic description of this variation.

My research also takes as an a priori assumption that language perpetually changes. Despite the relative stability of some sociolinguistic variables, the English language itself is always changing — as all languages do — and the potential exists that new variants of even these stable sociolinguistic variables may emerge and change the rules that govern the variable system (Tagliamonte 2006a, Ch. 1). In the following section I will discuss three ways in which language changes over time and space.

2.2 Changes through time and space

How do we account for similar linguistic features or similar linguistic changes between clearly different varieties of the same language? For example, Australian English and Texas English are in one sense very similar. Along
with London English, Dublin English, Vancouver English and Johannesburg English they share the large English lexicon and most of its grammatical rules, as well as the sense that some forms, like *ain’t* are less formal than other forms that mean the same thing, like *is not*. But Australian English is clearly different from Texas English; there are many differing features between the two varieties at each level of grammar and in the lexicon. So, when both Australian and Texas English show a similar increase of an innovative variant, or if they have the same incidence of a linguistic feature that is otherwise divergent from the rest of the English language diaspora, how do linguists account for it? In order to answer this question it is necessary to classify the ways in which language changes.

Labov (2007) discusses two ways in which linguistic change is propagated to new speakers: transmission and diffusion. Transmission is the unbroken sequence of native-like acquisition by children that results in a faithfully reproduced linguistic system, with community-wide linguistic change propelled forward via adolescent incrementation. Diffusion across communities, on the other hand, is linguistic change that results from adult-to-adult interaction and is thus the acquisition of new linguistic structures after the critical period. The consequence of diffusion, but not transmission, is the weakening of original patterns and the loss of structural features. When testing whether diffusion has occurred across communities one diagnostic for confirming diffusion is establishing 1) that one community showed the change first and 2) that the same community has a more structurally complex variable system. For example, if I hypothesize that a change in the quotative system diffused from Toronto to Cape Breton, good evidence to support this hypothesis would be that the change occurred first in Toronto, and that the constraints governing the quotative system there are more complex than those governing it in Cape Breton. For example, Tagliamonte and Denis (2014) examined quotative verbs in Toronto and several smaller communities in Southern Ontario, the authors found evidence suggesting that incipient *be like* was borrowed into the smaller communities from Toronto without borrowing the full set of internal linguistic constraints for the Toronto quotative system.

Toronto, Canada’s most populous city, is the vanguard of linguistic innovation in Canada; a new (or newly popular) feature may not be first attested in Toronto, but generally it must first take root in the city before it diffuses to other Canadian communities. This is in line with observations that densely-populated, politically, economically, and socially prominent urban centres, which generally have a lot of dialect/language contact, are incubators for language change, and that the direction of influence is usually from urban centres to more rural areas (see, for instance, Chambers and Trudgill 1998, 178; Trudgill 1974b; 1983; Bynon 1977, 214; McMahon 1994, 229; Weinreich, Labov, and Herzog 1968, 153-5; Callary 1975, 66–72; Hernández Campoy 2000; Britain 2002, 82). Industrial Cape Breton is as far as you can travel east from Toronto and still hear people speaking what linguists classify as Canadian English. This puts Cape Breton on the easternmost periphery of Canadian English, and thus it is a prime place to test the nature of diffusion as it occurs in a Canadian context. Chapter 6 will test the conclusions made by Tagliamonte and Denis (2014) that *be like* was diffused as a fixed expression, like a lexeme, in Canada and will show that this lexical-like diffusion occurred over great distances across both geographic and ostensible linguistic barriers and perhaps without sustained large-scale inter-varietal adult-to-adult communication.

If two communities where diffusion may have occurred instead show identical constraints governing a vari-

---

1. It is also possible that diffusion can occur without an observable loss of complexity. For example, if a feature is already maximally simple before it diffuses. In this case additional evidence is needed to confirm that diffusion has occurred.

2. Any further east and you arrive in Newfoundland which, though politically part of Canada, has traditional Englishes that have always been considered as *linguistically* distinct from Canadian English. For example, Boberg (2010, 26) writes "Indeed, that Newfoundland English is to be considered at all in a study of Canadian English is a result of purely political, not linguistic factors; linguistically, traditional Newfoundland English has more in common with the southwestern English and southeastern Irish varieties from which it historically derived than with mainland Canadian English."
able system, assuming that those constraints are complex enough that their simplification via diffusion would be expected, it is more likely that the change in question was not diffused from one community to the other, but that they both inherited the variable system from a period of the language that pre-dates the geographic or social divergence of the two communities. In other words, the variable system progresses through time via transmission in both communities. This is called parallel change. For changes that pre-date the origin of Canadian English, it has been observed that Toronto is not always at the forefront of change — instead the nature and timing of the founding of Canadian communities may play a more important role in determining how and when changes occur (for example deontic modality and stative possession as reported in Tagliamonte and Denis 2014). Thus it may be Cape Breton’s unique mix of founding populations, and the timing of the planting of the English language in the community (see Chapter 3), rather than its status as the final frontier of Canadian linguistic diffusion, that may determine how certain types of changes progress.

Assessing genetic relationships by comparing variable systems is referred to as “comparative sociolinguistics” — a method whereby the researcher aims to “discover usage patterns in the relative frequency of occurrence or co-occurrence of structures, rather than simply in the existence or grammaticality of those structures” (Poplack and Tagliamonte 2001, 91). As Durham (2014, 60) summarizes: “If two varieties present the same conditioning and direction of effects when examined using statistical analysis in the form of a multivariate analysis, then it is highly likely they come from the same source variety.”

Table 2.1: Accounting for similar changes across varieties.

<table>
<thead>
<tr>
<th>Change</th>
<th>Consequence</th>
<th>Requirement</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Variant $\alpha$ of variable $\nu$ diffuses from focal variety $A$ to peripheral variety $B$</td>
<td>Constraints for $\nu B$ are less complex than $\nu A$</td>
<td>Face-to-face communication by adults</td>
<td>Diffusion</td>
</tr>
<tr>
<td>2. Variety $A$ and variety $B$ are on parallel paths for variable $\nu$</td>
<td>Constraints for $\nu A$ and $\nu B$ are similar, though rates of variant $\alpha$ may vary</td>
<td>Varieties $A$ and $B$ are sisters, or have similar founding inputs</td>
<td>Transmission</td>
</tr>
<tr>
<td>3. Variant $\alpha$ of variable $\nu$ develops independently in variety $A$ and variety $B$</td>
<td>Variety $A$ and variety $B$ meet same structural conditions for variant $\alpha$ to arise</td>
<td>Varieties $A$ and $B$ are historically unrelated, but share similar grammatical structures for $\nu$</td>
<td>Drift</td>
</tr>
</tbody>
</table>

$\alpha$ = Variant  
$\nu$ = Variable

It is also possible for two varieties to show a similar change that is neither diffused, nor inherited from an earlier state of the language before the two varieties diverged. This similarity is the result of language drift. Following Sapir (1921), we can classify some changes as occurring simply due to the evolution of the grammatical system: “language moves down time in a current of its own making. It has drift” (Sapir 1921, 5). For example, the Canadian Shift is a change in the pronunciation of the TRAP, DRESS, and KIT phonemes in Canadian English following the merger of the low back vowels. Roeder and Gardner (2013) have argued that the merger of the low back vowels triggered a structural reconfiguration of the vowel phoneme system which resulted in predictable movement of TRAP, DRESS, and KIT. They also argue that this movement will occur in any North American varieties with the same vowel phoneme configuration.3 This is language drift. Language drift occurs when a common trigger already present in a previous state of the language is actuated resulting in change. This may occur at different times in different communities — for instance the Canadian Shift drift-

3. Though see, for instance, Durian (2012) and Bigham (2010) for alternative scenarios.
type actuation occurred first in southeastern Ontario (Clarke, Elms, and Youssef 1995) and British Columbia (Esling and Warkentyne 1993) before other parts of the country. As Roeder and Gardner (2013) have argued, for this type of change to occur in multiple communities, the communities need not be in contact, but rather they must share a similar structural configuration in the grammar that meets the criteria for the change to be triggered. The structural similarity may result from completely different dialect trajectories, and unlike parallel change, there need not be an unbroken succession of transmission of the change to the two communities from some previous common variety. Language drift is likely in situations where what looks like parallel change is occurring in two communities, but our knowledge of the socio-demographic history of the varieties in question suggests that transmission from a common source is highly unlikely.

2.3 Peripherality and diachrony

Cape Breton, at the eastern periphery of the Canadian English speech community, gives us a unique window into an earlier form of Canadian English. Tagliamonte (2013a, 14) calls peripheral English varieties linguistic “woolly mammoths” because, like examining woolly mammoths frozen in glaciers, examining peripheral English varieties allows linguists to gain insight into a completely different time — namely an earlier time in the history of the language. This is because linguistic changes often originate in central or “focal” areas and diffuse towards the periphery, or because changes occurring in a language often move at a slower pace through time in isolated rural communities with small, tight-knit populations than in high-contact urban communities with large, dynamic populations (Bonfante 1947; Coulmas 2008, 579; Chambers and Trudgill 1980, 109; 19(119,151),(585,235)(121,297),(598,382)(125,747),(606,834) 11; Trudgill 2011, 8). Cape Breton is likely to be one of the last eastward communities to receive diffused changes from Southern Ontario; its history of physical, social, and economic isolation, plus its relatively small size, suggests parallel changes ought to be at an earlier stage than in Toronto; and finally, with respect to language drift, if Cape Breton English changes slowly, triggered changes reliant on particular configurations of the grammar may occur at later times compared to more dynamic communities where the grammar evolves and meets triggering conditions more swiftly. Therefore, examining how linguistic changes (whatever their type) are occurring now in Cape Breton provides insight into how changes occurred in less peripheral areas of Canada in the past. Tagliamonte (2013a, 15) argues that peripheral areas “provide prime evidence about an earlier stage (or ancestor) of a language and play a key role in reconstructing earlier stages of a language’s development. Thus, critical evidence for determining the antecedents of a variety can be found in remote, inaccessible or otherwise isolated areas.” Industrial Cape Breton English is an excellent candidate for providing this type of evidence. Not only can we observe changes occurring through time, and study their nature, but by looking at especially the oldest speakers in Cape Breton we gain insight into earlier stages in Inland Canadian English. This permits us to probe hypotheses about diachronic change and earlier states of English grammar explored in studies using data from more mainstream varieties.

A confounding characteristic of peripheral communities to the potential observation of eastward linguistic diffusion in Canada is that, according to Schilling-Estes (2003), in some isolated English-speaking communities the use of highly-marked local features increases alongside increased exposure to mainstream varieties — often as a way to reinforce local identity and uniqueness. This means that on the one hand there is rapid adoption of non-salient new features, while on the other hand there is a resistance to or decreasing use of saliently non-local features. This is not particularly different from the garnering of covert prestige in the face of changes from

---

4. Woolly mammoths too survived thousands of years longer on peripheral islands in the Arctic Ocean and Bering Sea after they became extinct on mainland Eurasia (Veltre et al. 2004; Vartanyan, Garutt, and Sher 1993). These peripheral islands are dubbed the species’s “terminal refugia” (Lister and Stuart 2008, 615).
above (Trudgill 1974b, and elsewhere). Several studies in Newfoundland, for example, have found that young St. John’s speakers confronted with increasing exposure to mainland Canadian English are readopting traditional Newfoundland English speech features in order to show Newfoundland pride, or are embedding these traditional features in new or non-traditional ways in nascent social hierarchies (Childs et al. 2010). For example, stopping of /t/ and /d/ and -s marking on non-third-person-singular verbs encode localness among all urban females and are increasing in use, but for male speakers the features encode both localness and covert prestige so are increasingly becoming class-stratified (Childs et al. 2010). Labov’s (1963) original study of Martha’s Vineyard found that variability in the use of local or non-local (i.e., mainstream) diphthong pronunciations depended on whether island residents viewed speakers of mainstream varieties (i.e., summer tourists) positively or negatively. The extent to which diffusion of linguistic features towards Canada’s eastern periphery is helped or hindered by factors such as local pride or the evaluation of non-local speakers or speech must be taken into account when interpreting the results of this study.

2.4 Layering versus functional distinctions

For the four variables introduced in Section 1.2, there is an overriding question concerning the nature of each change: are innovative forms replacing older forms, or is the system evolving so that specific forms fulfill certain syntactic or semantic functions?

Layering describes part of the process of grammaticalization whereby nonce or formerly lexical expressions take on grammatical meanings that overlap with already existing functional forms (see, for instance, Hopper 1991, 22; Hopper and Traugott 2003, 125). Grammaticalizing forms may begin in restricted contexts, but gradually expand to all contexts where the forms over which they have been layered are used. Over time the innovative form may increase in frequency of use and the older form may become obsolete. To take expressions of future temporal reference as an example, the English lexical verb go in present progressive tense (e.g., be going) plus the preposition to has taken on the grammatical meaning of future temporal reference. The lexical meaning is roughly ‘being on the way to do something,’ e.g., I’m going to eat lobsters ‘I’m currently en route to a place where I will eat lobsters.’ The eating of lobsters will occur sometime after the utterance time, so it is easy to see how this expression eventually took on the grammatical meaning of future temporal reference. I’m going to eat lobsters can therefore also mean ‘I will eat lobsters sometime in the future.’ If be going to is layering over an older future temporal reference system it is available as an optional expression of this grammatical meaning in all contexts.

An alternative scenario is when a grammatical variant is introduced into the system and the new form and the older form become specialized for specific syntactic, semantic, or pragmatic functions (Hopper and Traugott 2003, 125). In other words, rather than one form taking over, each form finds its niche within the grammar. This co-existence may persist, but it may also be ephemeral as eventually one form expands in use, layering over, and ultimately replacing the other.5 In this scenario, for example, be going to may be used exclusively for future actions, while will is used for future states.

For each of the four variables, the question is whether innovative variants are layering over pre-existing variants or whether the variable system is becoming specialized. In order to explore this question, it will be imperative to carefully examine the grammatical constraints underlying the use of each variant. As Cape Breton

5. See Wallenberg (2013) who suggests specialization only occurs (rather than layering) when speakers attribute contrastive meanings (denotational or otherwise) to competing forms. The author also suggests stable variables result from competing forms being mapped to a continuous dimension of specialization, and that stability may be fleeting diachronically.
English likely represents an earlier state of Canadian English, it may be possible to capture grammaticalization of innovative forms among its oldest speakers.

2.5 Age, Sex and Social Status

In Western society linguistic variation is highly correlated to social stratification. Often, optionality in language use both reflects social divisions and helps to reinforce them. Time after time, variationist studies find that men and women, old people and young people, and high status individuals and low status individuals differ in their use of linguistic variables. For example, women and high-status individuals generally are found to have higher frequencies of certain variants (which come to be labelled “prestigious” or “standard”), while men and low-status individuals generally are found to have higher frequencies of other variants (which in turn are labelled “non-prestigious” or “non-standard”). Young people also generally have higher rates of diachronically newer variants than older speakers because of incrementation (Labov 2007, 346). These generalizations, however, are based on observations made in fairly stable communities, and Cape Breton has throughout its history been fairly unstable.\(^6\)

In the lifetime of the oldest residents alive in Cape Breton right now, Industrial Cape Breton has gone from the most prosperous industrial area in Eastern Canada to one of the most economically depressed areas in Canada. The three big industries that supported the region’s economy: steel; coal; and cod-fishing, no longer exist\(^7\) and with them went the social structures through which people understood their place in the community. These changes in the way in which Cape Breton society is stratified could have substantial effects on the linguistic use in the community. After a similar loss of industry in Pittsburgh, another steel town, formerly working-class dialect features were re-interpreted by speakers as indexing “localness” (Johnstone 2010). Therefore class stratifying variables became variables that stratify speakers by local affiliation, disrupting the former trajectory of incrementation. The same re-interpretation could be occurring in Cape Breton.\(^8\) If true, there ought to be observable class stratification for linguistic variables among older speakers but no or a different kind of stratification among younger speakers.

Linguistic innovation is brought about by language/dialect contact or by changes that occur within the community. This is Labov’s (2001) well-known dichotomy between change from above and change from below. In both cases young women are almost always the vanguards of change (Labov 2001, 266 and elsewhere). Innovative variants of the four variables may be increasing in use in Industrial Cape Breton, but can we expect that young women will lead the change? Between 1996 and 2011 the community lost more than 15 percent of its population due (mostly) to outmigration — a greater loss than any other Canadian census division (Statistics Canada 2012). The median age in the community in 2011 rose to 47.5\(^9\) because the biggest age cohort in the community (those born between 1945 and 1960) is aging, and also because after 1980, the age group leaving the community the fastest are those younger than 30. Migration to the community has been by mostly retirees (Environmental Design and Management Ltd. 2008). Between 2006 and 2011 the number of residents younger than 65 shrank by 6.8 percent while the number of residents over 65 grew by 4.4 percent. How does language change proceed in a community in which old people are replacing and outpacing the young and where the greatest contact with other dialects is by or has been by the oldest speakers in the community? Does the progression of

---

6. These generalizations may not hold up in minority and bilingual communities either, perhaps also due to instability. See Kang and Nagy (2016) for an overview of recent research.
7. Fishing of other species (especially lobster) does however persist.
8. Gardner (2010b) and Gardner (2013b) have suggested this to be the case with reference to phonological variables.
9. Canada’s median age in 2011 was 40.6 and Nova Scotia’s median age in 2011 was 43.7 (Statistics Canada 2012).
linguistic change in the community follow the regular pattern? And what happens to ongoing changes when the normal processes of incrementation are broken? My dissertation aims to address these questions.

Many variationist studies of Canadian Englishes do not test differences between social status, perhaps because, as Chambers (2010, p. 2 and elsewhere) says, “Canadians are not only highly urbanized but also overwhelmingly middle-class, to an extent that can scarcely be comprehended by outsiders.” This suggests there would be little sociolinguistic stratification based on social status in a Canadian context. That said, D’Arcy and Tagliamonte (2010) did find significant differences between occupation types for the probability of restrictive relative pronouns in Toronto English; they also found significant differences between education levels. Using the same corpus of Toronto English, Tagliamonte and Denis (2010) found that education was also a significant factor for the use of conservative general extenders, and Tagliamonte and D’Arcy (2007a) found the same to be significant for changes in deontic modality. Occupation was not reported, either because it was not tested or because it was found to be non-significant, suggesting that education rather than occupation may be a better marker of social status as it affects sociolinguistic stratification in Toronto English. Other quantitative variationist studies of Atlantic Canadian dialects (e.g., Childs et al. 2010; Roeder and Gardner 2013) have not included or reported on social status indices in their analyses.

As will be discussed in Chapter 4, operationalizing social status within this study will be tricky given the history of the community and the makeup of the data used; however, it will play a major role in understanding the results.

2.6 Summary

In this section I first presented the variationist sociolinguistic framework I will employ in my analysis of the three morphosyntactic and one discourse/pragmatic variables. I presented several sociolinguistic issues. It is my hope that my research will add an insightful voice to the ongoing discussions in the field about these issues. I outlined three ways in which linguistic changes propagate among new speakers — transmission, diffusion and drift — and presented criteria for assessing whether an observed pattern represents one of these three. I laid out the expectations and implications associated with conducting research in peripheral dialect areas and discussed how current changes to Cape Breton English likely mirror changes that occurred in Inland Canadian English in the past.

All variationist sociolinguistic research ought to examine internal linguistic and external social conditioning when considering variation between competing forms. In the above section I have presented several relevant overriding questions with respect to both internal and external constraints. For example, can the variation for these four variables be explained by the specialization of variants, or are variants layered over the entire functional domain? Or, what is the relationship between language change, sex, and age when the expected demographic trajectory is interrupted?

In the following sections I will touch on these issues and attempt to answer these questions as I present both the history of Cape Breton English and four case studies of current variation and change.
Chapter 3

Cape Breton English

It would be wrong to think of Cape Breton English as being a homogeneous entity. Both outsiders and Cape Breton residents view Cape Breton English as a collection of different dialects that intermingle on the island. In the broadest sense, we can divide Cape Breton English into four main groups: Industrial Cape Breton English (the focus of this research), Western Cape Breton English, Acadian Cape Breton English and Indigenous Cape Breton English. Industrial Cape Breton English — itself heterogeneous — is the Cape Breton English spoken in the Cape Breton Regional Municipality, the most populous part of Cape Breton, and the region whose settlement and growth resulted from its military and industrial steel-making, coal mining, and port operations. Western Cape Breton English is the Cape Breton English spoken in the rural areas of Richmond, Inverness and Victoria counties that were first settled by (mainly) Gaelic-speaking or Gaelic-English bilingual Scottish and Irish settlers. This is the dialect of Cape Breton English I assume most commentators are thinking about when they call Cape Breton a recessive Scottish English enclave. Presently Gaelic is moribund in these communities, though Celtic — heritage and culture — including music, dance and storytelling — are still vibrant and celebrated. Acadian Cape Breton English is spoken by the descendants of Acadian French-speaking settlers in the towns and immediate vicinities of Arichat and Chéticamp. Acadian French is still spoken (or understood) by many speakers in this community and Acadian heritage and culture is still alive and celebrated in these areas. Indigenous Cape Breton English is spoken in the five reserves of the Mi’kmaq First Nation on the island: Eskasoni, Membertou, Wagmatcook, Waycobah and Chapel Island. The Mi’kmaq language is still spoken and understood by many speakers in these communities and Mi’maq heritage and culture is still alive and celebrated. Unlike Gaelic, both French and the Mi’kmaq language currently enjoy institutional support through some public primary and secondary school education.

During Industrial Cape Breton’s industrialization, many residents from Richmond, Inverness and Victoria counties moved to Cape Breton County to work in the steel plant and coal mines, as did immigrants from Europe, the Middle East and the Caribbean. The Membertou reserve is also located within the boundaries of the Cape Breton Regional Municipality. Students from the reserve have attended primary and secondary school in the city, while adults have worked and shopped there too. In the last 15 years ties between Membertou and the surrounding Cape Breton Regional Municipality have strengthened, and the two are increasingly integrated. Thus, Western Cape Breton English, Acadian Cape Breton English, and Indigenous Cape Breton English are part of the mix of dialects that have led to contemporary Industrial Cape Breton English and also represent the most immediate dialects with which Industrial Cape Breton English has been in contact.

1. Celtic is the catchall word Cape Bretoners use to mean Scottish and Irish. These groups intermingled in the history of Cape Breton. See Section 3.1
3.1 The history of Cape Breton and its English

The following section traces the history of the planting of the English language in Cape Breton. Understanding this history is important for several reasons. Perhaps most importantly, the founding population of a community can have long-ranging effects on a local language variety. This is because generally the first effective settlers in an area set the community’s standards and norms for social behaviour, including the type/variety of language spoken. All those who follow must conform to the patterns of social behaviour that the first settlers establish. Labov (2001, 503–4) points out that Zelinksy’s (1992) social doctrine of “first effective settlement” — whereby the influence of new groups entering an established community is limited and the original group determines the cultural pattern of newcomers, even if the newcomers are more numerous — is consistent with the linguistic patterning of American cities like New York, Philadelphia, Boston and Chicago. Chambers (inter alia 2004a) and others have made the same observation for Canadian English-speaking communities.

Below I will trace the history of the English language in Cape Breton with a variationist sociolinguist’s eye in order to determine to what extent these claims of Cape Breton English’s “Celtic-ness” is true or even likely. For three of the four variables — future temporal reference, stative possession, deontic modality— it has been observed in other Canadian contexts that the timing of settlement and the nature of the input dialects correlates with differences in the variable system (e.g., Tagliamonte and Denis 2014). Therefore, these variables will act as test sites for the conclusions about the possible Scottish, Irish or other influence on Cape Breton speech that come about while examining the dialect’s history.

3.1.1 Pre-Industrial Cape Breton (late 1700s–late 1800s)

The first inhabitants of Cape Breton Island were the ancestors of the Mi’kmaq (Micmac) people, who spoke — and still speak, though use of the language is recessive — a language (also Mi’kmaq) that is part of the Eastern Algonquian language subfamily. John Cabot is credited with being the first European to discover Cape Breton, in 1497, though the province of Newfoundland and Labrador contests this and instead claims Newfoundland was the island Cabot found. A very small Portuguese settlement is known to have existed briefly, likely on the island’s northern coast, in the 1520s, though this settlement did not persist. Other Scottish and then French settlements also arose in the early 1600s, though none lasted and the island remained vacant of Europeans in the last half of the 17th century.

Following Queen Anne’s War (1702–1713) and the signing of the Treaty of Utrecht, France ceded its claims to the Hudson Bay area, Newfoundland and Acadia (present day mainland Nova Scotia, New Brunswick and northern Maine) to Britain; however, it retained several islands in the Gulf of St. Lawrence, including Cape Breton. Quickly France established a garrison on the eastern tip of the island, which it called Louisbourg, and began exploiting, albeit on a small scale, the vast mineral deposits on the island. Several very small fishing and trading communities also sprung up across the island, some by French-speaking Acadians who moved to Cape Breton from the British-controlled mainland. Louisbourg, with its ice-free harbour and (eventual) large fortress, became a naval and commercial hub. It also attracted a handful of Catholic Irish-Newfoundland fishing families, who settled outside the fortress walls and later further inland. In 1745 and again in 1758 New England naval forces sacked Louisbourg and each time razed the other small communities around the island. Privateering of

---

2. This observation was independently formulated as the “first-past-the-post” principle by Sankoff (1980) and as the “founder effect” by Mufwene (1996).
3. The same has been found for other variables, for example the merry-marry merger in southern Québec (Baxter 2010).
4. This section is mostly a synthesis of the very well-detailed history of Cape Breton Island presented by Morgan (2008) and the history of Cape Breton Scottish and Irish immigrants presented by Kennedy (2002) and A. A. MacKenzie (1999).
the island was also very prevalent. The Treaty of Paris in 1763, which ended the Seven Years War, put Cape Breton for the first time under British control. The French were evicted from Cape Breton and the British navy held Louisbourg. Even though the French were evicted, the few Irish-Newfoundlander fishing families were allowed to stay — but no new settlers joined them. A British order that same year (1763) forbade the granting of land in Cape Breton and effectively cut the island off from immigration.

This British ban on land grants is particularly important to the sociolinguistic history of Cape Breton because during the French and Indian War (the North American theatre of the Seven Years War) the British forcibly deported the French-speaking Acadians (who had from 1713 until 1755 been allowed to remain in the British-controlled territory of Acadia) and following the war, Britain encouraged northeastern New Englanders, as well as foreign Protestants, to come to Acadia and take over the Acadians’ farmland. The New England Planters, as they were called, were the first effective settlers in most communities in modern-day Nova Scotia and New Brunswick and their dialect served as the input for Canadian Maritime English (see, for example, Keifte and Kay-Raining Bird 2010; Algeo et al. 2001, 425; Roeder and Gardner 2013). However, these New England Planters were not sanctioned to settle in Cape Breton — the island thus had a fundamentally different input population from almost every other community in the Maritimes.

Following the American War for Independence in the late 1770s and early 1780s, approximately 30,000 United Empire Loyalists were evacuated to Nova Scotia, mainly from ports in New York and northeastern New England. During this same period the Scottish Highland Clearances led to several ship-loads of Gaelic-speaking Hebridean Scots immigrating to northwestern Nova Scotia. These two streams of immigration led to the creation of several new towns in Nova Scotia; however, at the time Cape Breton was still officially closed to immigration. United Empire Loyalists from western Pennsylvania and upstate New York, who by the 1770s spoke differently from their coastal contemporaries, were evacuated not through eastern coastal cities, but overland and into what at the time was referred to as Upper Canada — i.e., Southern Ontario. These inland United Empire Loyalists were the first effective English-speaking settlers in this region and are credited (by Chambers 2010 and others) with the setting the foundation for what would come to be known as Canadian English. The historical input variety for Toronto, then, is also quite different from the Maritimes and also Cape Breton. Loyalists from Western New England migrated northwards to southern Lower Canada (Québec) and the Ottawa Valley, again creating English-speaking communities with fundamentally different input dialects than in the Maritimes and Cape Breton.

As United Empire Loyalists streamed into the present-day provinces of Ontario and Québec, fearing that all the good land along the St. Lawrence would soon disappear, Québec governor General Federick Haldimand (among others) petitioned to have the functionally vacant Cape Breton opened up for settlement by Loyalists fleeing from nearby New England.

The first boat of 140 Loyalist settlers arrived from Québec, having travelled there the year prior from New England.

---

5. Some of these Acadians managed to flee mainland Nova Scotia and join the one small Cape Breton Acadian community near Arichat. Others fled and set up a small fishing village in northwestern Cape Breton, which they called Chéticamp.

6. The importance of Halifax should also be mentioned. It is the largest city in the Maritimes and the capital of the colony and then province of Nova Scotia. It began as Britain’s answer to Louisbourg during the first-half of the 1700s; the city is dominated by the garrison of the Halifax Citadel. The city has been without question the most important urban centre in the Maritimes since its founding. Its peopling was more diverse than other Maritime communities and its development was more top-down — the city was completely planned-out before the first boat of builders touched shore — so its original inhabitants were British officials, officers, engineers and soldiers, as well as hired builders of varied backgrounds. The city has set macro-regional standards for many things including language (Edwards and Jacobsen 1987). Past surveys of Canadian dialects that have commented on Maritime English (especially Keifte and Kay-Raining Bird 2010; Boberg 2010) have framed the divergent features of Halifax English as being different from Inland Canadian English, whereas they frame the divergent features of other Maritime dialects (Lunenburg English, Liverpool English, Cape Breton English, and African–Nova Scotian English) as being different from Halifax English.

7. A few small not-officially sanctioned settlements did occur in Western Cape Breton, though they were by Irish-Newfoundlanders or by Highland Scots who had previously landed in northwestern mainland Nova Scotia or Prince Edward Island and moved north.
York. Their first destination was Louisbourg, but when the settlers saw how unfit conditions were there (e.g., most of the viable building timber in the area had long since been felled, and the few Irish families living there had already cannibalized most of the remaining buildings), they decided to move on to St. Peters and then eventually Baddeck. The first boat to go to Spanish River (renamed Sydney) was the Blenheim which sailed from Portsmouth, England in October 1784. It contained 129 settlers composed of Loyalists, disbanded British (Loyalist) troops, and English immigrants. The Blenheim too arrived at Louisbourg, but after only five days decided to head up the coast and find somewhere else to settle, finally landing in Sydney. These two boat landings were significant for two reasons. First, they brought to Cape Breton the people who would become the island’s first successful English-speaking settlers: relatively well-off Loyalist families headed by well-connected, educated men (often former British officers) from the eastern seaboard of the now United States\(^8\) or from Britain; and second, the events of the boats’ arrival established the relationship between two English dialects: that of the Loyalists and that of the poorer, and increasingly marginalized Irish settlers — a dichotomy that would shape the character of Cape Breton English to the present.

This social dichotomy was again reinforced in the early 1790s when a boatload of starving Irish convicts wrecked near Main-á-Dieu. The population of Sydney at the time was only around 100 people. The residents had little choice but to allow the convicts to integrate into the community, albeit as servants or low-status workers, though this nearly starved the community to death.

Most of the Loyalists arrived in Cape Breton by 1793 — in total they only numbered about 500. They tended to remain in their original areas of settlement: Sydney, Baddeck, near Middle River, Port Hood, Hillsborough, and the Strait of Canso region. According to Morgan (2008, 112) the Loyalists, who uniquely exerted political control of the island only married each other "keeping their distinctness alive throughout the 19th century." More Irish-Newfoundlander settlers came to Cape Breton in the following years, but mostly to fish or work in the coal mines opening up in Sydney Mines and Low Point/Lingan. Gaelic-speaking Scottish Highlanders also began immigrating to the island, though mostly to its western half.\(^{8}\)

The Loyalists played a key role in the settlement of Cape Breton between 1784 and 1800. County capitals at Sydney, Port Hood and Baddeck, settlements of prime importance in the island’s history, were founded by them. They held key offices there and greatly influenced the early political and social development of the island, (Morgan 1975, 18)

But economic, political, cultural, and linguistic landscape of Cape Breton changed in the first half of the 1800s when the island received a massive influx of Gaelic-speaking Scottish settlers.

The fundamental character of Cape Breton Island was established with the settlement of the Highland Scots. Their great numbers caused a major change in population of Cape Breton and ultimately set the tone for the island. (Morgan 2008, 112).

Between 1802 and 1850 over 50,000 Scots (mostly coastal Highlanders) came to Cape Breton, joining the already established Hebridean Scots communities. The migration began in 1802 when 415 Scots landed in Sydney Harbour, liked the area, and sent word back to relatives and friends who were still aiming to leave Scotland because of the political, social, economic, and agricultural changes occurring in that county. Cape Breton was also the nearest point to Scotland with available arable land and therefore the cheapest place to reach. These newer Scottish settlers, especially those who arrived closer to 1850, were not as wealthy as the those who had emigrated from Scotland earlier. The new poorer settlers often took less desirable lands inland from the waterways, where

\(^8\) This included, for instance, the former mayor of New York
the wealthier earlier immigrants had settled, and were thus referred to as “Backlanders” (Davey and MacKinnon 2016, 7). Families followed families, with people of one area or island in Scotland settling next to their former neighbours. This segregation by settlement tended to keep people of opposing religious views apart. Morgan (2008) says the Irish inhabitants merged culturally with those Scots who were Catholic and that this blended group became quite successful.

Though the Scots-Irish were relatively few in number, they wielded substantial economic and social power by extending credit to the cash-poor settlers. Moreover, their ability to speak sophisticated English — unlike the Gaels, Acadians and Mi’kmaq — greatly enhanced their influence, both locally and in the wider world, (Morgan 2008, 112).

The historical experience of the Irish, along with their greater familiarity with English, made it logical that they would help provide a bridge between their Scots cousins and the outside world, and if this bridge was not always completely amicable or altruistic, without it the Highland adjustment to their life would probably have been more difficult, and it is perhaps not too much to suggest that the Irish bridge, by shielding the Scots, helped prolong Gaelic and its cultural milieu, (Toner 1999, viii)

It is not possible to know the exact linguistic background of the more than 50,000 Scots who migrated to Cape Breton in the 1800s; however, it is presumed that most were native Gaelic speakers. The Nova Scotian and Canadian governments did not include Gaelic as a language option on a census until 1931; however, historic accounts (newspaper articles, memoirs, etc.) by islanders and visitors indicate that the first generation of Gaels spoke Gaelic and that the language persisted as either a first or second language for several generations — despite a determined effort by the Nova Scotia government to stamp out the language in favour of English as a form of assimilation.

From a sociolinguistic perspective this is important, because it resulted in stratification by language and dialect. Gaelic speaking rural Backlanders traded with Gaelic-speaking/bilingual Scots-Irish, and the Scots-Irish English/bilingual speakers — whose English was the locally low-status Scots-Irish English dialect — traded with English-speaking Loyalists. Its no wonder that gradually the term “Celtic” became more commonly used in Cape Breton English to refer to anyone or anything with Irish or Scottish heritage, as the two cultures, from the point of view of the Loyalists, were confounded.

<table>
<thead>
<tr>
<th>Table 3.1: Linguistic stratification of Cape Breton, early 1800s.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Status</strong></td>
</tr>
<tr>
<td>Loyalists</td>
</tr>
<tr>
<td>Scotish-Irish</td>
</tr>
<tr>
<td>Low Status</td>
</tr>
</tbody>
</table>

Table 3.1 schematizes the quasi-diglossic language context in Cape Breton in the early 1800s described by Morgan (2008). At the top of the social continuum were the high-status Sydney Loyalists, who spoke a mixture of 19th century British and American Seaboard English. These two forms of English would have themselves been stratified socially. At the opposite end of the continuum are the Gaelic-speaking Scots, whose English was
heavily Gaelic-influenced L2 English. In between these two extremes were bilingual Scots-Irish speakers, who spoke both Gaelic and a Scottish/Irish-influenced English.

It is important to remember that Cape Breton did not exist in a vacuum, its isolation was not total. Trade from ports in Sydney, Arichat, Baddeck, and the Canso Strait area brought Cape Breton speakers into regular contact with merchants and sailors from mainland Nova Scotia (Bluenosers), the United States, Upper and Lower Canada (modern-day Ontario and Quebec), and Europe. Following the publication in Boston of Baddeck, And That Sort of Thing by Charles Dudley Warner in 1874, rural Cape Breton became a popular tourist destination, especially by affluent New Englanders (Yankees). As Tennyson (1986, xiv) points out, a common theme in the observations of late nineteenth century observers was the growing American influence on the island. In 1867, Sir John Bourinot claimed that “the Yankee element...is becoming very prevalent in Cape Breton,” noting that Americans owned “not only many of the coal and gold mines, but the principal stage-routes and the telegraph lines,” (Bourinot 1868, 196). Lady Aberdeen, who visited Cape Breton in 1897 with her husband, the governor general, commented in her journal on “the hordes of American tourists who come to the island every summer,” (Aberdeen 1960, 274). Like Warner, Lady Aberdeen was somewhat patronizing to the rural, mostly Gaelic-speaking residents of rural Cape Breton, saying, “nowhere have we met simpler, gentler, more hearty folk than these,”(Aberdeen 1960, 274). Lady Aberdeen was “sadly humiliated” that she could not converse with some rural women who only spoke Gaelic, but was enthused by her ability to communicate with children. “Of course, the younger generation speak English, regular public schools having been formed here in 1865,” (Aberdeen 1960, 274).

Social stratification on the island was somewhat dynamic. Sydney, the focal population area of the region, was mired in fights for power and prestige. While elsewhere in Cape Breton the social hierarchy centred around the local clergy or merchant, in Sydney the social hierarchy was much more complex. Among the upper classes there was intense bickering and rivalry for governance between the first group of Loyalist settlers and new arrivals . Morgan (2008, 185) reports that even after annexation to mainland Nova Scotia in 1820 (which resulted in part from this infighting), “Sydney society bore the signs of a colonial capital, with the Loyalist families, former government officials, and retired military officers at the top.”

As Kennedy (2002, 30) notes while describing the death of Gaelic in Nova Scotia, “however important bi-culturalism or multiculturalism might eventually become to Canadians, it is impossible to trace it to any pro-multicultural philosophy that was in place in Britain when Canada was founded.” The English language in 19th and early 20th century Nova Scotia was indisputably the language of the most powerful sector of the population — a sector backed up by an institutional infrastructure that increasingly consolidated its position of authority (Kennedy 2002, 47). The social primacy of English over Gaelic was thus part of Nova Scotia’s British colonial inheritance. For these reason there were strong social and market pressures for children of non-English-first-language parents to speak only English. These pressures were exacerbated by policies and practices of Nova Scotia educators to delegitimize Gaelic as a viable modern language. Writes Kennedy (2002):

Rather than dealing with the main obstacle to education in Gaelic communities…educators helped to perpetuate the myth that Gaelic was an inherently “backward” language unsuited to advanced intellectual function and incapable of fulfilling any institutional role. Moreover, they propounded the belief that Gaelic’s mere existence presented a formidable obstacle to the education and advancement of Gaels...The continued exclusion of Gaelic from education and such normal domains of use

9. Based on 19th-century literary representations, e.g., McKinnon (1852)
would have multifarious effects and was probably the single most important cause of the unusually rapid decline of the language in Nova Scotia in the 20th century. Because the domains of exclusion for lower or non-status languages are most commonly the domains of wealth, power, and prestige, it is understandable that speakers of those languages should be inclined to make the practical decision to learn the language that appears to give them best access to those desired ends. The urban environment proved particularly hostile to the socialization of children in Gaelic. The Gaelic language and Gaelic traditions were much less likely to be passed on from one generation to the next in urban areas than they were in rural districts, (Kennedy 2002, 49,58,78).

Unlike other colonial Englishes, in which an English speaking group is in contact with usually-larger indigenous population, the primary contact context in 19th-century Cape Breton was between Loyalist English-speaking colonists or Yankee visitors and a larger group of Gaelic-speaking later settlers. Prior to the advent of major coal mining and steelmaking operations in Industrial Cape Breton in the late 19th century, Gaelic-speaking rural Scots and Loyalist-origin colonists lived in geographically separate parts of Cape Breton Island. Their interactions were largely transactional and often mediated by bilingual speakers (see Table 3.1 on p. 20). As Morgan (2008, 112) points out, intermarriage between Loyalist-origin colonists and other settlers was rare, and as Kennedy (2002) details, there was institutional pressure during this time in rural schools for children to use only English. Sustained contact between Gaelic and English speakers likely did not occur until eastern Cape Breton’s industrial operations drew rural residents to the towns surrounding Sydney and Glace Bay (along with new immigrants) at the turn of the 20th century.

In summary, pre-19th-century European settlement in Cape Breton is moot with respect to the transplanting of English to the island. The founding English-speaking population to Cape Breton was Loyalist, though Irish-English speakers and L2-English Gaelic, Acadian French and Mi’kmaq speakers were a significant proportion of the island’s population. Given the founder effect theory, contemporary Cape Breton speech will match more closely other Loyalist varieties (like mainstream Canadian English) rather than varieties from the United Kingdom, though Cape Breton English may show additional transfer from other native or L2 varieties. Furthermore, from the time of the planting of English in Cape Breton, the Loyalist variety of English was established as the overtly prestigious variety as the Loyalists and their descendants held positions of authority within the community and wealthy Yankee tourists were the most common interlopers.

### 3.1.2 Cape Breton English in the 20th-Century

Despite government development initiatives to turn the economy around, the towns in industrial Cape Breton have known more economic uncertainty than prosperity through most of [the 20th] century, (Gardiner-Barber 2002, 401).

While contemporary accounts of Cape Breton discuss a culturally-Celtic population unified by island-pride and contempt towards Halifax or the mainland generally, Cape Breton in the mid-1800s was a collection of distinct Mi’kmaq, Acadian, Scottish, Scots-Irish, and Loyalist communities. In the areas where these peoples came together — around ports like North Sydney and mines like those in Sydney Mines — there was clear social stratification. Morgan (2008, 99) argues that the racial, historical and economic interests of the island’s settlers were too divergent at the time to permit a unified sense of island identity or “Cape Breton-ness.”

---

10. Although there was non-negligible contact between the Mi’kmaq people and colonial English speakers in Cape Breton, for the present discussion, I focus on contact between colonial English speakers and Gaelic settlers.

11. The use of literary dialect tropes by McKinnon (1852), the clergyman and newspaper editor of Loyalist descent credited as being Cape Breton’s first novelist, is an excellent representation of this.
Coal production in the eastern communities of Cape Breton County, especially Glace Bay, ramped up in the mid to late 1800s. The expanding mining operations attracted many residents from the already overpopulated rural areas and weighted the island towards its industrial base.\textsuperscript{12} The opening of the steel mill in Sydney, and its offshoot industries, drew further immigrants not just from the British Isles, but also from eastern and southern Europe, the Middle East, the West Indies, and Newfoundland. The area was the most dynamic growth zone in Atlantic Canada up to World War I. This boom period also saw historic labour organizing and major strikes (which continued through the 1920s), and represents one of the most interesting and militant periods of Canadian labour history (Gardiner-Barber 2002). Gardiner-Barber (2002) suggests that a strong sense of local identity was brought about during the hardship of early settlement of Cape Breton Island and then later reinforced by working-class families suffering through the poor working and living conditions and major industrial strikes of the late 19th and early 20th century.

Of note is that while many of the coalminers and steelworkers were of Loyalist descent, the vast majority were either new immigrants, their descendants, or Scottish/Irish, Acadian, or Mi’kmaq workers from other island communities. Language in the mines or at the plant was thus a conglomeration of L2 English, Scottish/Irish English, other non-standard English, or often not English at all. The same can be said for the neighbourhoods that built up around the Sydney Steel Plant and the coal mines — especially Whitney Pier, whose development was recognized by the Government of Canada in 2015 as being “an event of national historic significance.”

Between 1899 and 1930, the neighbourhood of Whitney Pier emerged with the largest concentration of ethnocultural groups in the Maritimes, making it one of the most distinctive immigrant communities in Canada during the 20th century, paralleled in diversity only by much larger immigrant centres in Montréal, Toronto, and Winnipeg, (Parks Canada 2015).

More than 15 ethnic and religious groups established themselves in Whitney Pier and to varying degrees persist in Whitney Pier or elsewhere in Industrial Cape Breton today (Parks Canada 2015; Pottie 2015).

During both WWI and WWII, Cape Breton’s industrial operations, ports, and proximity to Europe meant that it was an important strategic Canadian asset, especially as a staging ground for merchant convoys.\textsuperscript{13} During both conflicts Cape Breton was thus a “hive of wartime activity,” (Morgan 2009, 52). Cape Bretoners were also ready volunteers as soldiers. The wars brought Cape Bretoners, both soldiers and those back home, into constant contact with speakers of other varieties of English (Canadian, American, British, etc.).

The early 20th-century industrial period was also the period of rapid decline in the use of Gaelic on the island. The lure of employment or education opportunities in bigger English-speaking communities drew away the children of the rural Gaelic-speaking farmers, hindering the perpetuation of the language \textit{in situ} (Dunn 1953, 133).

The young people who grew up in the pioneering communities unwittingly tended to associate the Gaelic language which they heard at that time with the incessant toil, hardship, and scarcity peculiar to primitive conditions. When they went to the city, the universal language was English, while Gaelic was unknown; and the standard of living there was inconceivably superior to what they had known. Hence Gaelic came to be considered the language of poverty and ignorance and was therefore despised, while English was regarded as the language of refinement and culture and therefore cherished, (Dunn 1953, 145).

The 1931 census, the first to ask about Gaelic, found that 24,000 Nova Scotians (about 6 percent of the Nova Scotia population), mostly living in Cape Breton, were native Gaelic speakers — though most could also speak

\textsuperscript{12} The history of Cape Breton Island in the twentieth century is synthesized from Morgan (2009). See also Gardner (2010b).

\textsuperscript{13} Cape Breton’s, and especially Sydney’s, military history is extensively documented by Tennyson and Sarty (2000)
One speaker from the Post-Industrial corpus who was born in 1921 and grew up on rural Boularderie Island, says her L1 Gaelic parents only used the language with visiting relatives, or while attending church.

(5) **You’d see them laughing, carrying on. Oh, Great Uncle John come over and he be laughing and laughing. And then say “Did you hear this one?” And then they’d talk and talk, you know, you know, you might get a word here and there, that would connect with it, that’s it.**

(6) **In church, we had the first service in English and the second service in Gaelic. And of course, I could sing all the Gaelic hymns, you know, with them. I didn’t know what I was saying, but I could go along with them, you know?**

**Interviewer** I guess it would be no different than, like, Catholic people in Latin?

**Speaker** Yeah. Same idea.

Until 1921 children were forbidden to speak Gaelic in Cape Breton schools — a practice carried over from the 19th century, and so most children were accustomed to speaking with each other only in English, even in rural areas. Explains the same speaker:

(7) **Well, they all spoke Gaelic, pretty much like I did. You know what I mean? They, they were the same way. Their parents spoke Gaelic mostly, but, I don’t know if they spoke it in the house or not, but they, ah, we, as children, didn’t speak to each other in Gaelic. It was always just English.**

**Interviewer** And your parents always spoke to you in -

**Speaker** Oh yes, always in English, yes.

In 1921 Gaelic was permitted as an optional subject in the school curriculum, but lack of qualified instructors meant this had a negligible effect on preserving Gaelic as a vital living language on the island (Dunn 1953, 157). Coal production in Industrial Cape Breton peaked in the 1940s; in 1965 the private company operating the coal mines and steel plant in Industrial Cape Breton announced the mines had only 15 years of production left, and that the company would be pulling out of the community within months. Responding to public outcry the federal government created a crown corporation to run the mines, and the provincial government took over the steel making operations. Just prior to these changes in industry, the community’s population peaked at 131,507 (in 1961, see Figure 1.1 on p. 7), but since this change of hands, and through the gradual decline of these industries, the population has fallen below 1931 levels. The Lingan Colliery closed in 1992, followed by the Phalen Colliery in 1999 and the Prince Colliery in 2001. At the same time, the provincial government decided to dismantle and sell the steel plant. During this period (1996 to 2001) the CBRM experienced the largest population percentage decline of any Canadian census division. The population drop gutted Industrial Cape Breton’s working and middle classes. Members of both classes were required to find work in new sectors or move off-island for work or further education. From 1975 to 2002 the average rate of unemployment on the island was 18.2 percent, compared to 11.3 percent on mainland Nova Scotia. In 1993 the island’s unemployment rate reached 29.5 percent, so that almost one in three people in the community did not have a job (Locke and Tomblin 2003, 8; Morgan 2008, 232). Gardiner-Barber (2002) calls residents’ continued loyalty to the community in the face of social and economic hardship “the culture of making do.”

Gardiner-Barber (2002) explains that each generation of Cape Bretoners has maintained a commitment to the local cultural identity and a commitment to community loyalty despite the grim economic realities of living on the island.

“Making-do” involves expressing community loyalty in the face of adversity. It also involves setting a united front against outsiders (including people from Halifax, especially government) and a strong
commitment to kinship, relatives, and immediate family. To the extent that the loyalties to one’s kin and community avoid, obscure, or even deny the negative qualities of life in Cape Breton families and communities, we may speak of the loyalties as ideological, (Gardiner-Barber 2002, 400-401).

This ideology of loyalty may result in the maintenance of non-standard vernacular forms despite the hegemony of “standard” (i.e., Standard American or Standard Canadian) varieties. Residents can employ local features as part of their performative and agentive co-construction of local identity14 — an identity they are expected to create in order to belong to the local community and “present a united front against outsiders,” as Gardiner-Barber (2002) notes above. Westhaver (1996, 94-95) suggests iconic local linguistic forms “act as communicative markers that differentiate the people of Cape Breton from those living elsewhere.” She argues iconic Cape Breton linguistic forms “provide us [i.e., Cape Bretoners] with markers of self-identification, they close social gaps, and they strengthen friendships.” And while these features may reflect the “genuinely good” characteristics of Cape Bretoners, they may also “serve as a form of caricature of the culture. Nevertheless, Cape Bretoners cling tenaciously to [iconic features] as a means of identification and a way of preserving [their] down-home personae.” Paradoxically, the “down-home persona” that is constructed out of such hardship is one that reinforces it, as the archetypal Cape Bretoner is considered, particularly off-island, to be jobless, drunk and uneducated.

While the most hostile views of Cape Breton might conjure images of a scenic paradise inhabited by idiots, even would-be sympathetic observers are given to contrasting our ‘hospitable’ nature, humorous character, and archetypical ‘party-animal’ status with an assumption on our dependence on welfare, Unemployment Insurance and other government handouts, (Robertson 1991, 8).

To “sound Cape Breton,” especially towards non-islanders, is an ideological stance incongruous with sounding and thus being successful (though this may be changing, see below). To embrace and actively index a Cape Breton identity is to embrace and actively index an identity that marks a person as incapable of being sober, smart or fiscally responsible. Cape Bretoners who want to be successful may thus actively avoid sounding local because they do not buy into the associated ideology or do not wish to take on the identity that speaking this way entails. For example, one speaker from the Post-Industrial corpus made the following comment not in relation to sounding young, or girly, or unintelligent, but in relation to sounding like a Cape Bretoner

While the most hostile views of Cape Breton might conjure images of a scenic paradise inhabited by idiots, even would-be sympathetic observers are given to contrasting our ‘hospitable’ nature, humorous character, and archetypical ‘party-animal’ status with an assumption on our dependence on welfare, Unemployment Insurance and other government handouts, (Robertson 1991, 8).

To “sound Cape Breton,” especially towards non-islanders, is an ideological stance incongruous with sounding and thus being successful (though this may be changing, see below). To embrace and actively index a Cape Breton identity is to embrace and actively index an identity that marks a person as incapable of being sober, smart or fiscally responsible. Cape Bretoners who want to be successful may thus actively avoid sounding local because they do not buy into the associated ideology or do not wish to take on the identity that speaking this way entails. For example, one speaker from the Post-Industrial corpus made the following comment not in relation to sounding young, or girly, or unintelligent, but in relation to sounding like a Cape Bretoner

What it means to “sound Cajun” in Louisiana according to Dubois and Horvath (inter alia 1998) may prove insightful in looking at these alternative pressures to both sound and not sound local. As the status of Cajuns changed with the economic development of their communities and the rise of the education level of community members, being Cajun became something to be proud of. A “Cajun Renaissance” occurred and the meaning of sounding Cajun changed as Cajun food and music became popular, drawing tourists to Cajun communities. Now, while grandfathers sound Cajun and speak French because it was the vernacular and language of their childhood, and middle-aged Cajuns sound more like white English-speaking Louisiana city-dwellers because they grew up having negative associations with sounding Cajun, grandsons are taking up Cajun Vernacular English features both out of pride and because it is financially beneficial if they work in the Cajun tourism industry. Thus both ideological and market pressures work in concert to promote sounding Cajun among young speakers.

14. Potentially any non-standard form could be available for this type of local identity construction as Cape Breton English speakers confound “local” and “non-standard” language features. See Section 5.3, for example.
Boudreau and White (2004) make similar observations in Chéticamp, the small Acadian French-speaking community on the shores of western Cape Breton. They remark that the increase in “heritage tourism” has reinforced the distinctive features of the community’s Acadian dialect. Here again the market pressures speakers to employ local or traditional forms. Boudreau and White (2004) also point out that the increase in tourism has brought the local dialect into greater contact with other varieties of French and thus is also exerting greater standardizing pressure. For English-speaking Nova Scotia the promotion of heritage tourism is often referred to as “tartanism.” Under the premiership of Angus L. MacDonald (1933-1954), the province of Nova Scotia (Latin for New Scotland), “purposely and politically embraced and celebrated Scottish lineage as its common ancestry,” (Rolls 1996, 79; Smith-Piovesan 1998, 31–32). While this did boost tourism, especially for Cape Breton, it also reintroduced the 19th-century image of Cape Breton as a quaint, traditional and unspoiled island, and reinforced the association between being from Cape Breton and being backward, naive, uncultured, unsophisticated, or simple. While the increase of tourism has opened up Cape Breton communities to interactions with outsiders—a phenomenon ostensibly challenging to linguistic conservatism (e.g., Milroy 1980), market pressures may also persuade residents to maintain quaint, traditional or unspoiled social practices—including ways of speaking. The indexical field for locally-evaluated speech features may grow to include not just negative qualities, but also positive qualities like commercial success, entrepreneurial spirit, or hospitableness.15

Deindustrialization and the demographic changes that were its consequence occurred in parallel with Cape Breton’s “cultural renaissance” from the 1960s onward. Cape Breton’s Magazine, first published in 1971, retold the stories of islanders in their own, unedited words. The success of the magazine led its editors to start Breton Books, the island’s first nationally-distributed trade publisher, which published new local literature and reissued classic Cape Breton music compositions. A few years earlier, in 1966, Elizabeth and Harry Boardmore inaugurated the first theatrical studies program at what is now called Cape Breton University, and promoted the on-stage presentation of Cape Breton stories to Cape Bretoners. Influenced by the Boardmores and burgeoning musical culture, groups of young people produced a series of revues that blended elements of the island’s Scottish, Irish, French (Acadian) and working-class cultures. Writes Morgan in his 2009 history of Cape Breton, the revues “played to packed and delighted local audiences who really understood the humour, the irony, and felt the pride in what could be made of their own story.” (Morgan 2009, 211). Out of these revues grew The Rise and Follies of Cape Breton Island, later re-envisioned as the Cape Breton Summertime Revue, which ran almost annually from 1977 to 1996. These massively popular revues drew much of their humour from the use—or misuse—of local language, explains both Morgan and Smith-Piovesan:

From hardboard bleachers the crowds laughed and cheered [at] scenes that featured Cape Breton music, clever socio-political satire, gut-busting humour, much of it unintelligible to off-islanders, (Morgan 2009, 212).

The Cape Breton accent is the focal point of much of the humour. In addition, one character in particular, Martin MacKinnon, is noted for his exaggerated mispronunciation and misuse of words. Double entendres appear quite often in the (Cape Breton Summertime Review) as stereotypically uneducated, simple Cape Bretoners poke fun at those from the mainland [i.e., mainland Nova Scotia] or the United States who are perceived as having more intelligence and worldly experience, (Smith-Piovesan 1998, 68).

The Follies and the Revue, as well as other iconic Cape Bretoners on stage and on television (e.g., the character

---

15. MacPherson and Corbin’s (1996, 241) survey of 70 Cape Bretoners about national and local identities revealed the most common view is that Cape Bretoners are “friendly, unemployed, hard-working party-animals, who are proud of Cape Breton.” Less often-cited qualities were “polite, generous, hospitable, happy, family-oriented, and big beer-drinkers.”
of General John Cabot Trail and musicians like the Rankin Family and Natalie McMaster) helped to reinforce the cultural stereotypes associated with being from Cape Breton and to enregister the local dialect. Two of the shows’ songs, Kenzie MacNeil’s *The Island* and Leon Dubinsky’s *We Rise Again*, have become unofficial anthems for the island. The catchphrases of the shows’ characters, like “Good dear, good!” and “What’s your father’s name?” have now come to index the community generally and can be found written on mugs, T-shirts, bumper stickers and hats at Cape Breton souvenir shops. Language forms like “Good dear, good!” now connote a Cape Breton identity, but the identity that they connote is an iconic or stereotypical identity. This stereotypical Cape Breton is filled with hardworking and hospitable unemployed, uneducated drunks; and for young speakers it is this Cape Breton this is exactly what is indexed when using traditional speech features.

Photograph 1: Cape Breton T-shirt featuring the phrase “Good dear, good!” Photographed by the author, July 4, 2017.

### 3.2 Previous Accounts of Cape Breton English

The first cataloguing of Cape Breton English speech features come from an article by W. D. Lighthall that appeared in the Toronto periodical *The Week* in 1889 (Lighthall 1889). Lighthall called Bluenoser speech “obviously almost pure Yankee” outside of the more British-sounding Halifax. The speech of Cape Breton and Pictou Counties’ Scottish descendants he labels “Acadian Scotch.” The salient features of Acadian Scotch include (summarized by Orkin (1970, 14-15)):

1. The *s* is often pronounced with the soft instead of the harsh should, as in ‘reserve,’ pronounced somewhat as if written ‘re-serve’.

2. The *u* is often pronounced as if *y* preceded it, as in ‘Jerusalem’ which you will hear pronounced as ‘Jeryusalem’ or as if *h* preceded, as in ‘pursue’ pronounced somewhat as if written ‘purshue’.

3. A common and most characteristic turn of speech consists of the use of ‘whatever,’ sometimes in the sense of ‘at any rate,’ as in the following sentences: ‘The crop is very good whatever.’ ‘Money may be plentiful, but the times are bad, whatever.’
4. So, too, there is a peculiar use of the word ‘altogether’ as meaning much the same as ‘very’ or ‘extremely.’ ‘He is a good preacher altogether,’ i.e., an eminently good preacher.

5. Among the people of Highland decent we find many peculiarities owing to the use of negatives in Gaelic idioms, which are foreign to our English tongue, such as ‘It is a long time since I did not see you.’

6. Among those whose knowledge of English is limited there is a great confusion in the use of pronouns. So marked and so common is this that it has given rise to the popular saying that a Highlander calls everything ‘she’ except his wife, who is always ‘he.’ The effect is sometimes ludicrous.

7. Again there is sometimes a singular transposition of prepositions. Thus a friend of mine inquired of a stranger whence he came. The answer was ‘From Cape Breton over.’

8. Another expression often heard here among housewives is apt to strike a stranger only. Bread when heavy is said to be ‘sad.’

As will be shown below, the later analyses of Cape Breton speech data show some of the same features persist (namely 1, 4, 5 6), while others, like sad bread, may have been lost to time.

The earliest systematic analysis of specifically Cape Breton English is Rowe’s (1968) dialect survey of the Lake Ainslie area of Inverness County, on the western side of the island. At the time of her survey, this area of Inverness County was considered to be the most Gaelic on the island, if not North America generally (Rowe 1968, 4; J. L. Campbell 1936, 129). Rowe’s (1968) intention was to probe the extent to which the use of regional speech features in the Lake Ainslie area reflected the area’s Gaelic heritage. Prior to Rowe’s work, there were four linguistic interviews from Cape Breton speakers collected by Henry Alexander for The Linguistic Atlas of the United States and Canada (Alexander 1940; 1939), though these records were not published (Dollinger 2008, 34). Rowe modelled her surveys after Alexander’s, so her work is comparable to reported findings from New England (Kurath et al. 1939–1943, etc.) and from similar studies in King’s and Lunenburg counties in mainland Nova Scotia (Wanamaker 1965; 1980; R. Wilson 1958). She included additional questions in order to make direct comparisons to available accounts of Scottish English/Scots (Grant and Murison 1931–1976; Wright 1898–1905; J. Wilson 1915; Jamieson 1808-1841).

Overall, Rowe found overwhelming evidence that the form of English used in Cape Breton’s Gaelic heartland was Loyalist, with very little influence from Scottish English or Scottish Gaelic. She found that the use of dialect-region-defining lexical items among both older and younger speakers in the Lake Ainslie area matched predominately those of Loyalist-settled mainland Nova Scotia and the Loyalists’ original lands, New England. This included words like johnny cake, darning needle and curds (p.30–31), which are (or were ostensibly) only common in these areas. There was strong use of terms considered Canadianisms, like chesterfield (‘couch’), eavestroughs (‘gutters’) and bureau (‘chest of drawers’) (p.31, 36), which are generally not or no longer found readily in either American or British English (e.g., Boberg 2010; Burkette 2001).

There was no evidence of Gaelic lexical transfer, aside from keadack! (a sheep call) used by one bilingual speaker (p. 32). Interestingly, some of Rowe’s informants indicated that other locals use the words poke, galluses, and bucket, which the informants believed erroneously were of Gaelic origin, though Rowe found no evidence of the words’ usage anyway. In fact, even keadack! may be kin to the sheep calls /ko-dejk/ and /ko-daek/ recorded by Kurath (1941, 28) in New England.

There were a number of non-standard verb forms (e.g., the use of the preterite for the past participle, or vice versa) and pronunciations (e.g., stopping of interdental fricatives) reported by Rowe (1968), though the extent to which these can be classified as “Gaelic transfer” or simply as English vernacular universals is not clear.16

16. Of course, the instance of some English vernacular universals (like the stopping of interdental fricatives) are also directly linked to
The three features Rowe (1968) points out as being regularly-used Scotticisms in the community are the use of *swole* as the preterite of *swell*, the *after*-perfect and the use of *would* with a progressive or future meaning.

(9) a. My hand swole up (‘My hand swelled up’). (Rowe 1968, 67)
   b. I am just after going to town. (‘I just got back’) (Rowe 1968, 80)
   c. When would she be leaving? (‘When is she leaving?’) (Rowe 1968, 80)

Wright (1898–1905, Vol. 5, p. 880) lists *swole* as found in both Inverness, Scotland, and in the rural Pennsylvania dialect represented in Lloyd’s (1900) *The Chronic Loafer* — so it is equally likely to be of Scottish or Loyalist origin (though whether the coastal Loyalists who travelled to Cape Breton shared this feature with their Pennsylvanian brethren, or whether similar Scottish migration to rural Pennsylvania brought *swole* to the area is also unknown). The *after*-perfect, on the other hand, is well-documented in Newfoundland English (Clarke 2010, Sec. 3.2.5; etc.), with a clear provenance in Irish and Highland Scottish settled areas of the island. It is also attested on nearby Prince Edward Island (Pratt 1988, 4) and in Ulster-Scots settled areas of the United States (Montgomery 2006, 34). The feature, a combination of *be* + *after* + the present participle is a direct calque from Irish/Scottish Gaelic (Ó Sé 2004). Finally, progressive *would* does not have a well-established connection to vernacular Scottish (or Irish) English or Irish/Scottish Gaelic. Rowe cites Wright’s (1898–1905, Vol. 6, pp. 493–495) claim that the use of *will* to express simple present where there is some element of conjecture is a Scotticism. Rowe extends this to posit that progressive *would* may also be Scottish.

In summary, with the exception of the *after*-perfect, Rowe (1968) finds little conclusive evidence of Scottish English or Gaelic influence in the English of speakers from the Lake Ainslie area.

Falk (1989; 1990) explores the use of non-standard linguistic features in the oral narratives published in *Down North* — a collection of interviews from *Cape Breton’s Magazine*. The majority of the non-standard features she identifies are attested and often frequent across vernacular varieties of English (e.g., non-standard *be* agreement in existential and coordinate clauses; demonstrative *them*; *a*-prefixing; preterite and past participle levelling for some verbs; *for to* in infinitival clauses, etc., see Kortmann and Szmrecsanyi 2004). However, three morphosyntactic features identified by Falk (1989) are notable: non-emphatic, non-contrastive topicalization (e.g., (10) a., b. below); subject relative pronoun deletion in existential clauses (e.g., (10) c., d.); and the use of gendered pronouns for third person singular inanimates (e.g., (10) e.).

(10) a. My cousin he made all kinds of knots.
   b. Kept putting a little water in her mouth Mrs. McLennan did.
   c. There was bad things [đəh] happened there.
   d. And there were hundreds [łuńu] used to come here.
   e. You wouldn’t have that in your pocket... I’d have to take one quarter out of her [= pocket].

(Falk 1989, extracted from Caplan 1980)

These morphosyntactic features could be used as strong evidence for Gaelic or Scottish/Irish-English substrate effects in Cape Breton English because they are features with either direct cognates in Scottish Gaelic/Irish or are prevalent among Scottish/Irish English speakers. However, they may not be useful in distinguishing Cape Breton English from other varieties of Canadian English. For example, topicalization of a syntactic element via movement to the leftward edge of a clause is found throughout English varieties, although Kallen (2013, 77) does indicate it is more frequent in Irish English and may involve structures that are not used elsewhere (see also Miller 2004, for attestations in Scots, and Shuken 1984 for Highland and Hebridean English). Furthermore, it is a feature of Celtic languages (Fife 2009, 13), and may be borrowed into or develop in varieties of English English speakers from specific areas or specific social groups within the British Isles (including Ireland and Scotland) and these variants were later transported to the English-speaking diaspora.

17. These same interviews make up part of the *Storyteller* corpus used in the present study and described in Chapter 4.
where Celtic languages are substrate (Hickey 2005, 246). However, the similar doubling of subjects and objects on the left periphery is a feature of colloquial varieties of Canadian French (e.g., Auger 1994; Nadasdi 2000; Beaulieu and Balcom 1998); and subject doubling has been attested in the English of Montréal Anglophones (Nagy, Blondeau, and Auger 2003).

(11) a. My friend Martine’s French.
   b. The Québécois they know how to party.
   c. My sister she’s a music teach in Joliet.

(Examples from Montréal Anglophones, Nagy, Blondeau, and Auger 2003, 80)

The existence of topicalization in Cape Breton English cannot be definitively linked solely to Scottish Gaelic/Irish, or Scottish/Irish English substrate effects, given the feature’s attestation elsewhere in Canada and in other varieties. The extent to which the feature is used among speakers and the specific contexts in which the feature is possible across varieties are unsettled; therefore, a comparative study in which Cape Breton English is included would be fruitful.

The second feature of note identified by Falk (1989) is the deletion of subject relative pronouns in existential clauses. Overall, the deletion of relativizers (like that or who/which) in subject clauses is rare or ungrammatical in mainstream varieties of English. Deletion is however attested in many non-mainstream varieties, including vernacular Scottish English, though also several vernacular varieties of British and North American English and Anglophone Caribbean creoles. In these varieties deletion in existential constructions is more common than in other subject constructions (e.g., Rickford 2011; Miller 1993; Harris 1993; Tagliamonte 2002; Wolfram and Schilling-Estes 2016).

Finally, the use of gendered pronouns for inanimates is a feature of Scottish Gaelic, a language with grammatical gender. Shaw (1997) claims that the use of gendered pronouns for inanimates is a clear example of a Gaelic calque in Cape Breton English. While this is possible, it may also be a calque from Acadian French — which also employs grammatical gender — or borrowed from any of a number of languages introduced to the island during Cape Breton’s industrialization (e.g., Polish, Arabic). The use of gendered pronouns is also common in vernacular varieties of English outside of Cape Breton (see Wagner 2003, for a detailed analysis), including vernacular Inland Canadian English. This is especially true of her, which is often used in vernacular Canadian English when referring to abstract or difficult to identify referents, like ‘the ongoing state of affairs’ or ‘the situation at hand’ — as in the lexicalized phrases steady as she goes, or give’er (Clarke 2010, 85; McLeans 2013). The iconic Cape Breton greeting, How’s she going, b’y? employs one of these gendered pronouns in this way. Therefore, it is difficult to discern whether the use of he/him/his/she/her/hers instead of it/its is in fact a borrowed Gaelic feature in Cape Breton English, or instead a feature of vernacular Canadian English more generally.

Beyond the three constructions identified above, Falk (1989) also points to the use of since + negated preterite, rather than since + perfect forms, as in (12). Again, Shaw (1997) argues this feature stems directly from Gaelic influence.

(12) a. Well, it’s been 45 years since I didn’t make that. [vs. ... 45 years since I haven’t made that.] 18 (example from Falk 1989)
   b. It is a long time since I did not see you. (example from Shaw 1997).

Nearby Newfoundland English and other vernacular North American varieties are known to employ the preterite form frequently in several contexts where Standard English calls for perfect forms (Clarke 2010, 79).

18. I am a native Cape Breton English speaker and while the original quotation from Caplan (1980) is unnatural-sounding to me, Falk’s (1989) repair sounds equally unnatural. My intuition of unnaturalness stems not from the use of since + negative perfect forms, but rather the use of since with negative constructions. Of note, I was born in 1983 and grew up in Cape Breton’s urban-industrial area. My intuition may reflect the convergence of Cape Breton English with mainland norms.
Furthermore, in Newfoundland English and in other English varieties, never can have both a durative and a punctual meaning (Clarke 2010, 98). It is therefore possible that Cape Breton English shares these two features with Newfoundland and other varieties — leading to utterances like (12) above — rather than the phenomenon Falk (1989) and Shaw (1997) cite being a unique Cape Breton Gaelic transfer.

Finally, Falk (1989; 1990) attests several "non-standard" adverbial and adjectival intensifiers in Cape Breton English (too, awful, pretty, quite, right, altogether). The Canadian English intensifier system is particularly susceptible to rapid change and includes a large number of both historic and innovative variants. A further investigation of the patterns of intensifier use in Cape Breton English may prove fruitful for further exploring the connection between Cape Breton and mainstream Canadian English.

The remaining features that Shaw (1997) identifies as being the result of Gaelic influence on Cape Breton speech are phonetic/phonological: the use of [s] for canonical /z/ intervocally and word-finally (e.g., in words like reserve, busy, and position); pre-aspiration of voiceless consonants; and palatalization of velars before and after /o/.

The Atlas of North American English (Labov, Ash, and Boberg 2006, 216–224) included a handful of speakers from Nova Scotia, including two from Cape Breton. Of note, Cape Breton and Newfoundland diverged from the rest of Canada by showing a relatively retracted /a/ and a relatively high /i/, while the Atlantic region as a whole appeared to not participate in the Canadian Shift (lowering/retracting of /æ, e, ı/) and have a fronted /aa/. Speakers from Nova Scotia also showed distinctive tensing of /æ/ before nasals and /g/.

In fact, aside from these features, Cape Breton English was found to align with the rest of Canada along several dimensions; for example, not having a split short-a system (whereby BATH and TRAP words are distinct) and a merger of /lw/ and /w/, which distinguish several North American varieties from British and Scottish varieties. Cape Breton also participates in Canadian Raising in canonical environments.

Keifte and Kay-Raining Bird (2010, 68–69) find similar phonetic/phonological features in their survey of Cape Breton speech. They point out that Cape Breton speech shows similar features to Newfoundland, Halifax and Ontario. In fact, they place Cape Breton speech on a continuum between Halifax and Newfoundland, and attribute the similarity between Cape Breton and Newfoundland to a similar Celtic history on the two islands, and the direct immigration of Irish descendants via Newfoundland to Cape Breton. They attribute the similarities between Cape Breton English and Inland Canadian English (e.g., the diphthongization of the GOAT and GOOSE vowels) to market forces.

Because the economic situation in Cape Breton is quite depressed, we see the largest seasonal migration between the island and the Western Provinces and it is entirely possible that this has influenced local dialects in general, (Keifte and Kay-Raining Bird 2010, 68).

The authors note the following features of Cape Breton English: variable absence of the KIT/DRESS distinction; raising of TRAP before /d/ and /g/; variable monophthongization of FACE; variable monophthongization or centralization of GOAT; centralization of GOOSE; the use of a slit-fricative or glottal stop for /t/; low-level stopping of interdental fricatives; and Canadian Raising, though with the phonetic merger of MOUTH and GOAT in pre-voiceless contexts.

---

19. Subsequent research across Atlantic Canada, including Cape Breton, has found substantial evidence of the presence of the Canadian Shift, (e.g., Roeder and Gardner (2013), Sadlier-Brown and Tamminga (2008), and Hofmann (2014)).

20. The authors do not discuss interprovincial variation for this feature. See also Boberg (2010, Ch. 5.1) which reports on this pattern generally across Canada, and shows Maritime speakers and Ontario speakers behave similarly (relative to speakers from other Canadian provinces) in their raising of /æ/.

21. These features have also been observed in Cape Breton English by Parris (2009), Gardner (2013b; 2013a) and Roeder and Gardner (2013).
The authors report interspeaker variability in the merger of the LOT/THOUGHT vowels; however, Gardner (2009) showed that Cape Breton speakers do not show variability in the merger, but instead use the horizontal aspect of the low back vowel pronunciation stylistically, with more backed pronunciations in formal contexts, and more front pronunciations in informal contexts. Roeder and Gardner (2013) argued this was a reanalysis of two low-vowel systems inherited by Cape Breton English from different groups of Loyalist settlers.

Table 3.2: Principal Vowels of Cape Breton English — adapted from Keifte and Kay-Raining Bird (2010, 69–70).

<table>
<thead>
<tr>
<th>KIT</th>
<th>FLEECE</th>
<th>NEAR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>KIT</td>
<td>I, I</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>DRESS</td>
<td>I, I</td>
<td>NEAR</td>
<td>I</td>
</tr>
<tr>
<td>TRAP</td>
<td>PALM</td>
<td>A-a</td>
<td></td>
</tr>
<tr>
<td>LOT</td>
<td>THOUGHT</td>
<td>NORTH</td>
<td></td>
</tr>
<tr>
<td>STRUT</td>
<td>GOAT</td>
<td>O-o</td>
<td>O</td>
</tr>
<tr>
<td>FOOT</td>
<td>GOOSE</td>
<td>U-u</td>
<td>U</td>
</tr>
<tr>
<td>BATH</td>
<td>PRICE</td>
<td>A-A</td>
<td></td>
</tr>
<tr>
<td>CLOTH</td>
<td>VOICE</td>
<td>O-e</td>
<td></td>
</tr>
<tr>
<td>NURSE</td>
<td>MOUTH</td>
<td>O-U</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3: Principal Vowels of Standard Scottish English — adapted from Scobbie, Gordeeva, and Matthews (2006, Tables 2, 3).

<table>
<thead>
<tr>
<th>KIT</th>
<th>FLEECE</th>
<th>NEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIT</td>
<td>I-o</td>
<td>I-i</td>
</tr>
<tr>
<td>DRESS</td>
<td>E</td>
<td>E-e</td>
</tr>
<tr>
<td>TRAP</td>
<td>PALM</td>
<td>A-A</td>
</tr>
<tr>
<td>LOT</td>
<td>THOUGHT</td>
<td>O-o</td>
</tr>
<tr>
<td>STRUT</td>
<td>GOAT</td>
<td>O-O</td>
</tr>
<tr>
<td>FOOT</td>
<td>GOOSE</td>
<td>U-U</td>
</tr>
<tr>
<td>BATH</td>
<td>PRICE</td>
<td>A-A</td>
</tr>
<tr>
<td>CLOTH</td>
<td>VOICE</td>
<td>O-e</td>
</tr>
<tr>
<td>NURSE</td>
<td>MOUTH</td>
<td>O-U</td>
</tr>
</tbody>
</table>

It is important to point out that the vowel and consonant systems of Cape Breton English — as reported by Keifte and Kay-Raining Bird (2010) — are nearly identical to those reported for mainland Canada by Labov, Ash, and Boberg (2006) and Boberg (2010), perhaps not in the minutiae of phonetic implementation, but nearly entirely in regards to phonetic inventory and lexical incidence. This suggests minor regional phonetic variation, but ultimately the same underlying phonological system. On the other hand, the phonetic/phonological system of Cape Breton English is very different from that of conservative Scottish English. Scottish English does not differentiate between FOOT and GOOSE but does differentiate between /hw/ and /w/, and between NORTH and FORCE. Further, Cape Breton shows no evidence of the Scottish Vowel Lengthening Rule, the use of /x/, or a trilled /r/.  

There is one strong similarity between Cape Breton English and other Celtic Englishes. The use of a slit fricative [T] for word-final, utterance final /t/ or /d/ (Gardner 2013a; 2013b). This feature is found in nearby Newfoundland, and also in other English varieties with Irish English or Irish/Gaelic input. It is perhaps the

23. The IPA transcription and exact description of this sound is somewhat controversial, as discussed by Pandeli et al. (1997). The lack of a phonetic symbol for this sound may have led researchers of Irish English to mistakenly claim it only occurs in Ireland (e.g., Hickey 1986, 17; 1996, 224-225). I adopt the symbol advocated by Pandeli et al. (1997), and following Gardner’s (2013a) confirmation that the sound as produced by Cape Bretoners is an apico-alveolar slit fricative, refer to it as ‘slit-t’.
strongest evidence for Celtic influence on Cape Breton English; however, this influence is specifically Irish, not Scottish.

Finally, the recent publication, *Dictionary of Cape Breton English*, which contains just over 670 headwords of Cape Breton English words/phrases, includes 51 loanwords/calques in Cape Breton English from Irish/Scottish Gaelic. Table 3.4 summarizes these words by lexical domain. The majority of these words are used in English to describe preserved Gaelic cultural practices. These include words used when describing aspects of traditional Cape Breton Celtic music (e.g., *Ceòl Mor* ‘big music’), the names of supernatural beings in Cape Breton Celtic folktales (e.g., *bochdan* ‘evil spirit’), names for traditional foods (e.g., *isbean*, a type of sausage), and names for elements of traditional cloth making (e.g., *clò* ‘homespun woollen cloth’). All of these words, along with *loch* used in place names, and words for religious practices, household items, and farming, fishing, and nautical terms refer to specific referents inherited from Cape Breton’s Scottish and Irish immigrant culture. The exceptions to this are Gaelic terms used for insulting people, Gaelic terms used as greetings and interjections, and Gaelic terms used in nicknames. Most of the citations for each of these entries, though, come from Cape Breton literature and not from recorded speech.

The one grammatical feature in Table 3.4 is the use of the after-perfect, as described above. Whether this feature was an independent transfer from Scottish Gaelic to English in Cape Breton or brought to Cape Breton via English-speaking Scottish Highlanders or Irish Newfoundlanders is a question for future investigation. Of note, the after-perfect appears sporadically in the *Storyteller*; *Steelworker* and *Post-Industrial* corpora from speakers born from 1887 to 1960. No speaker born after 1960 employs the after-perfect; however, a systematic analysis of all the potential contexts in which the after-perfect could have occurred across the data was not performed, and must be left for future inquiries.

Table 3.4: Gaelic-origin borrowings/calques in Cape Breton English — based on Davey and MacKinnon (2016).

<table>
<thead>
<tr>
<th>Lexical Domain</th>
<th>Headwords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folklore and storytelling</td>
<td>7</td>
</tr>
<tr>
<td>Traditional music</td>
<td>6</td>
</tr>
<tr>
<td>Traditional foods</td>
<td>6</td>
</tr>
<tr>
<td>Used in nicknames</td>
<td>5</td>
</tr>
<tr>
<td>Insults or curses</td>
<td>5</td>
</tr>
<tr>
<td>Farming/fishing/nautical terms</td>
<td>5</td>
</tr>
<tr>
<td>Traditional cloth-making</td>
<td>4</td>
</tr>
<tr>
<td>Greetings and interjections</td>
<td>3</td>
</tr>
<tr>
<td>Religious practices</td>
<td>3</td>
</tr>
<tr>
<td>Household items</td>
<td>2</td>
</tr>
<tr>
<td>Used in place names</td>
<td>1</td>
</tr>
<tr>
<td>Weather</td>
<td>1</td>
</tr>
<tr>
<td>Mining terms</td>
<td>1</td>
</tr>
<tr>
<td><strong>Functional constructions</strong></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>51</strong></td>
</tr>
</tbody>
</table>

24. The greeting *ciad mile failte* ‘one hundred thousand welcomes’ retains currency in the community primarily as a tourism slogan.
25. Creative nicknaming practices in Cape Breton have been documented by the same authors elsewhere (Davey and MacKinnon 1996). In their analysis they do not list any pervasive Gaelic nicknames, although many of the English nicknames they do cite are for individuals and families with Scottish- or Irish-origin surnames (e.g., the Pickle Arse McLeans, the Red, Blue, and Ditty McDonalds, Biscuit Foot MacKinnon, among others. Davey and MacKinnon 1996, 206-207).
26. The data that makes up the *Storyteller* corpus is one of a number of corpora of Cape Breton speech/writing that Davey and MacKinnon (2016) draw from. Many of the citations that appear in the dictionary are from this data.
27. *Storyteller*: 14 instances; *Steelworker*: 35 instances; *Post-Industrial*: 5 instances.
One token from one of the oldest speakers in the Storyteller corpus is noteworthy. For this speaker, after is used in a perfect form, but not the canonical be+after+present participle. This speaker employs a blended perfect with both perfect have+past participle and after (see below). Beyond this token, this speaker employs a number of conservative vernacular forms (non-standard s-marking, preterite come, perfective done, among others). His blended perfect form may represent a broader use of after to mark perfect aspect in vernacular Atlantic Canadian English that has subsequently faded from use, or — as he was born and grew up during the waning years of Scottish Gaelic’s status as the most common Cape Breton home language — a period of increased idiosyncratic grammatical transfer from Gaelic to English.

Accounts of phonetic and phonological features in the community also come from folk linguistics. Da Mudder Tongue and its sequel Anudder Mudder (Gray 2006; 2007) are slang dictionaries for Industrial Cape Breton. The little photocopied and stapled books are sold at souvenir shops and local bookstores. Each entry includes approximations of phonetic transcription. A linguist reading Da Mudder Tongue can easily pick out linguistic features considered to be iconically “Cape Breton.” These include phonetic/phonological features like schwa elision (BUNGLO for ‘bungalow’, CUMP KNEE for ‘company’, etc.), vowel simplification (SANE for ‘saying’, TAL for ‘towel’, etc.), medial /t/ voicing (DIG NA DEE for ‘dignity’, BODDUM for bottom, etc.), dark /l/ deletion (AMOST for ‘almost’, AWRITE for ‘ alright’, etc.), yod deletion (ANNA RISM for ‘aneurism’, REGLUR for ‘regular’, etc.), interdental stopping (CAT LICK for ‘Catholic’, TOT for ‘thought’, etc.), /a3/—/o3/ neutralization (BILE for ‘boil’, BOITE for ‘bite’, BYE for ‘boy’, ICE TERS for ‘oysters’, etc.), /aw/—/aw/ merger before voiceless consonants (MOTE for ‘mouth’, TROUT for both ‘trout’ and ‘throat’, WIT OAT A DOTE for ‘without a doubt’, etc.), -ern metathesis (EASTRIN for ‘eastern’, PATRINS for ‘patterns’, etc.), sandhi /h/ insertion (HANG A HUEY for ‘hang a U-y’), and intrusive /i/ (WARSH, CHICAGO); morphosyntactic features like both YAWL and EWES as second person plural pronouns, some and right as adjectival intensifiers (SOME RITE NICE), alveolar -ing (YAKIN for ‘yakking’, TRINE for ‘trying’), third person singular female pronouns to reference a general state of affairs (BOOT NER for ‘booting her’, DRIVE NER for ‘driving her’, etc.), the use of ain’t, ME for ‘my’, non-standard simple past or past participle forms (NODE for ‘have known’ and NUDE for ‘known’, OWNID for owned, GOAD for ‘went’, etc.), non-standard comparative forms (GOODER and WORSTER), the addition of dynamic prepositions to to (and sometimes subsequent to deletion) as in “overt to the licker store” and “out da bunglo,” singular/plural mismatch with existential and deictic pronouns (“Dat wurnt da way it happened”), and the phonetic reduction of modal and semi-modal constructions (SHUDDA, CUPP, SHUDDENOV, KUDDENOV, GODDA, HADDA, etc.); and discourse-pragmatic features like the use of ME DEAR, BUDDY and BYE as discourse markers and HOW’S SHE GOAN? as a greeting. Clearly these features are not just perceived features of Cape Breton English. Many

28. Both mean to proceed quickly, or to work hard. “I was really busy at the call center today, I was boot ner and drive ner all day!”
are attested in other communities (e.g., nearly all appear in Newfoundland and Labrador English, see Clarke 2010) and many consist of English vernacular universals (especially alveolar -ing, Labov 1972, though see also Chambers 2004b; 2004a; Filppula, Klemola, and Paulasto 2009; Schreier 2009; Trudgill 2009). Many of the lexical items and nearly all of the idioms are not specific to Cape Breton, including: cheeses, pogey, dibs, frig, galoot, scoff.\footnote{While non-standard, these words can all be found in vernacular Inland Canadian or Newfoundland English, see Barber (2004; 2008) and Story, Kirwin, and Widdowson (1999)} your goose is cooked, tough as nails, the cat’s meow, talk a blue streak, loaded to the gills, son of a gun, it never rains but it pours, crazy as a bag of hammers,\footnote{All these idioms, and most others in Gray (2006; 2007) appear in Ayto (2009) or on the website UrbanDictionary, suggesting much wider use than just Cape Breton English} etc. It is telling that what is considered universally vernacular or just vernacular broadly coincides with what is considered iconically “Cape Breton.” Universal non-standard features are inextricable from local non-standard features, reinforcing the idea that to be “local” is to be non-standard. For this reason speakers of Cape Breton English may find it hard to disentangle what is broadly vernacular in Atlantic Canada, Canada, North America or the whole English speaking world, with what is vernacular because it is a specific local word, phrase, or pronunciation. The very existence of a slang dictionary of Cape Breton English and souvenirs decorated with Cape Breton catchphrases (like Photograph 2) entails that the features of the Cape Breton dialect, whether locally or universally vernacular, are part of the local discourse. Further, the commodification of these vernacular forms implies a process of local vernacular norm formation, or enregisterment, (Johnstone, Andrus, and Danielson 2006; Johnstone 2010) has occurred. In the last few years, use of local dialect features, especially off-island and in the media, has been both harshly criticized (M. Campbell 2003; J. Visser 2007) and applauded (Reid 2007; Ainslie 2007; Murphy 2008) suggesting that the dual nature of local vernacular forms, which index both local affiliation and local deficiency, is also part of the public discourse.

Photograph 2: Cape Breton T-shirt. T-shirts like this point to the commodification of Cape Breton English, especially in the context of tourism. Alongside local idioms like “Are you from the Bay bye?” are non-local vernacular idioms like “Holy Mackeral”, “Cat got yer tung?” and “Son of a Gun!”, or iconic vernacular Newfoundland English idioms like “Oh me nerves”, being categorized as iconically ‘Cape Breton’. Photographed by the author, April 28, 2017.

is local is particularly striking when compared with the results of Edwards and Jacobsen (1987). In a dialect perception study that involved 40 university students at St. Francis Xavier University (Nova Scotia) from Newfoundland, Cape Breton, mainland Nova Scotia and Massachusetts listening to and judging speakers from the same four locales reading an extract from John Steinbeck’s *Canary Row*, Cape Bretoners could only identify Cape Breton voices 20 percent of the time. The Cape Bretoners were 98 percent accurate in placing Massachusetts speakers, 98 percent accurate in placing mainland Nova Scotia speakers, and 40 percent accurate in placing Newfoundlander. The most common mistake by Cape Bretoners judging Cape Breton speakers was labelling them as mainland Nova Scotians — in other words, the Cape Breton speakers’ reference for a Cape Breton voice was somehow different (likely more non-standard) than the actual Cape Breton voices they were hearing. Cape Bretoners also frequently mistook Newfoundland voices as mainland Nova Scotian, suggesting Cape Bretoners’ reference for a Newfoundland voice was also different (again, likely more non-standard).

If Atlantic Canadian accents fall along a continuum of mainland Nova Scotia — Cape Breton — and Newfoundland (Keifte and Kay-Raining Bird 2010, 68 suggest this is the case), with mainland Nova Scotia being more standard and Newfoundland being less standard, and with place standing in for standardness, it is noteworthy then that judges from Massachusetts and mainland Nova Scotia most commonly misidentified Cape Breton voices as being from Newfoundland. Just as Cape Bretoners’ internal reference for a Cape Breton accent was more non-standard than an actual Cape Breton accent, outsiders assumed that an actual Cape Breton accent was more non-standard than it was. It is no wonder that Cape Breton t-shirt makers, or Cape Breton t-shirt buyers for that matter, consider any non-standard word or phrase to be licit as a local Cape Breton speech feature, whether it is local or not.

In summary, previous accounts of Cape Breton English have investigated the relationship between Cape Breton English and Scottish Gaelic/Scottish English. Try as they may, they do not find evidence of a dialect “more or less identical with the English of the Scottish Highlands” as Trudgill (2004) claims. In fact, there is strong evidence that Cape Breton English is simply a more vernacular form of Canadian English, with some
unique phonetic and lexical features. While some Scottish Gaelic/Scottish English transfer may have occurred (e.g., the after-perfect), this does not outweigh the preponderance of evidence linking Cape Breton English to other Loyalist-rooted varieties, and to a lesser extent Newfoundland English. The social milieu in which Cape Breton English exists is one in which Cape Breton English is judged to be non-standard (whether due to Celtic origins or not) both by locals and by others. For this reason any non-standard speech feature is judged also be a local Cape Breton feature. In the following chapters I will test the genetic connection between Cape Breton English, Inland Canadian English, and Scottish English by looking deeper than incidence to the rules underlying the variable grammar within the community. I will also show how, for some variables, the view that Cape Breton English is non-standard and that non-standard speech is local, has influenced the trajectory of certain variants.
Part II

Methods
Chapter 4

The Data

The data for this project is unprecedented in research on peripheral varieties of Canadian English. For this project four sources of data have been combined to create the *Corpus of Cape Breton English*. These data sources include the dialogue extracted from a mid-19th century Cape Breton author, the old stories of rural Cape Breton residents collected for a local oral history magazine, interviews with retired steelworkers about the workings and closure of the Sydney Steel Plant, and sociolinguistic interviews with especially young Cape Bretoners recorded by the author. From these four sources a total of 17,333 tokens from 358 individual speakers were extracted across the four studies detailed in the following chapters. This makes the *Corpus of Cape Breton English* not just a new but also a substantial tool for examining language change, its transmission, and its diffusion in a Canadian context. The tokens used for exploring stative possession, future temporal reference, and deontic modality are, in fact, the largest quantitative sociolinguistic datasets for these changes in progress in any variety of English to date and thus offer a significant view into the perpetuation and patterning of these changes. The dataset for quotatives is also the largest amassed in Canada outside of Toronto.

The four datasets are the *McKinnon Texts*, the *Storyteller* corpus, the *Steelworker* corpus and the *Corpus of Post-Industrial Cape Breton English*. Below each of these are described.

4.1 *The McKinnon Texts*

William C. McKinnon, a clergyman, newspaper editor, and descendant of United Empire Loyalists is credited as being Cape Breton’s first novelist (Morgan 2008, 190). In the 1850s McKinnon published several serialized novels that take place in Cape Breton: *St. Castine: a legend of Cape Breton; Frances, or, Pirate cove: a legend of Cape Breton; and The Midnight Murder*. These novels are made up almost entirely of dialogue, and it is from this dialogue that linguistic data has been extracted. For example, only the underlined portions of quotations in (15) were considered for extraction:

(15)  At this moment a Celt, Scot, or whatever other name is borned by the Americanized emigrant from the western isles of Scotland, entered the door — the rain dripping from his bonnet. “Foor, you please, te noight, sir,” said he to the landlady, on entering, in the broken English peculiar to them, “she’l give us a glace groc, you please?” The landlady poured it out in silence while the Yankee’s eye glistened, as if he had discovered new food for sport. (McKinnon 1852, 5)

While patterns of variation in the *McKinnon Texts* may align with modern spoken data, the *McKinnon Texts* are perhaps more helpful for establishing what variants existed in the community in the 1850s. Even though
Table 4.1: Tokens extracted from the McKinnon Texts.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quotative Verbs</td>
<td>—</td>
</tr>
<tr>
<td>Stative Possession</td>
<td>129</td>
</tr>
<tr>
<td>Deontic Modality</td>
<td>83</td>
</tr>
<tr>
<td>Future Temporal Reference</td>
<td>460</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>672</strong></td>
</tr>
</tbody>
</table>

literary works have their own set of conventions that are often markedly different from speech, if the McKinnon Texts show be going to as a marker of future temporal reference or robust have got use for stative possession, it would indicate that these variants were at least in the consciousness of early Cape Breton English speakers. Furthermore, the types of characters to whom McKinnon assigns particular variants is suggestive of the attitudes towards those variants. For example, the choice McKinnon makes to use the pronunciation spelling te noight in (15) above, suggests both that /æj/ centralization was present among at least some speakers in the community, and that those speakers were most likely an “Americanized emigrant from the western isles of Scotland.”

As the data from the McKinnon Texts is a literary representation of speech, and not speech itself, it will be analyzed qualitatively and independently of the speech data from the Storyteller, Steelworker, and Post-Industrial corpora. There are no quotations inside quotations in The Midnight Murder or the other words — so the McKinnon Text data offers little history with respect to the Cape Breton quotative system — but there are a number of tokens of the remaining three morphosyntactic variables. The 672 tokens extracted from the McKinnon Texts are not included in the 17,333 token count mentioned above.

Several studies have used 19th century Canadian literature to assess the origins and development of Canadian English morphosyntax. Among these studies are Dollinger’s (2008) analysis of modal auxiliaries (including modals expressing both deontic modality and future temporal reference) in early Canadian newspapers and letters, D’Arcy’s (2015) analysis of stative possessives in Victoria, B.C., newspapers, and Jankowski’s (2004) analyses of deontic modality and stative possession in early Canadian plays. To date, the works of McKinnon have not been analyzed within a variationist framework — nor have the works of any Cape Breton author. The McKinnon Texts data is therefore extremely useful, not just for providing insight into early patterns of variation in Cape Breton English, but also for assessing the similarities and differences between early Cape Breton and Inland Canadian English.

Not only do the MacKinnon Texts provide of a point of diachronic comparison, they also provide a point of synchronic comparison. In the early 1800s two other authors of very different backgrounds were writing similar stories in Nova Scotia. These two authors were Thomas Chandler Haliburton, the first bestselling author from

---

1. See, for example, Labov (1982, 20; 1994, 11) for an analysis of the non-optimal nature of historical text data.
2. McKinnon very much aped the "historical fiction"-style of the then popular Sir Walter Scott, which meant that he showcased alongside the high class protagonists a cast of colourful local characters employing vernacular speech represented orthographically, (G. Davies 1991).
3. The pronunciation of /æj/ with a central nucleus ([æj]) is a feature of several Scottish and Irish English varieties, which is often represented as ‘oy’ or ‘oi’ in literary pronunciation spelling. The use of a rounded and central nucleus for /æj/ and/or the neutralization of the /æj/ — /æj/ contrast is still present in conservative Prince Edward Island, Nova Scotia and Newfoundland English, as well as varieties on the eastern seaboard of the United States, (see Table 3.2, Keifte and Kay-Raining Bird (2010, 65), Pratt (1988), and Wolfram (2006), and the spelling pronunciations presented by Gray (2006; 2007)). Of course, the author could be representing Canadian Raising here as well.
4. Dollinger’s (2010) and Chambers’s (2008) analyses of representations of the low back vowels in early Canadian writing do analyze the works of Thomas Chandler Haliburton, a Nova Scotian author; however, these analyses do not comment on the use of any of these four variables in the current study.
what is now Canada, and Thomas McCulloch, a Scot living in Pictou County, Nova Scotia, whose *The Letters of Mephibosheth Stepsure* (1821–1822) includes letters written in the voices of Nova Scotia Loyalist and Scottish characters.

Haliburton, who was born in Windsor, Nova Scotia, to a second-generation New England Planter father and a New England Loyalist mother (R. A. Davies 2005, Ch. 2), relied heavily on literary dialect writing, as was the tradition in British literature in the 19th century, representing the speech of Nova Scotians and his protagonist Sam Slick, a Yankee, using non-standard grammatical constructions, regional words, and pronunciation spelling. Haliburton’s skill at using dialect writing as a literary device was particularly praised at the time according to biographer Cogswell (1976):

> There are many reasons why Haliburton’s greatest success was achieved as a writer… To the sense of proportion which he had acquired through his classical studies and his reading of 18th-century prose was added an ear sensitive to the rhythms and nuances of colloquial speech at the very time on the North American frontier when laxity of education was enabling that speech to run riot in a picturesque and racy manner not known in the English language since Elizabethan times, (Cogswell 1976).

> [Haliburton] did for the dialect of his time what Robert Burns did for Lowland Scots. The purist would say that both wrote a bastard language, but to most others their writing is a successful tour de force… Haliburton’s use of language added American to Lowland Scots on the list of English variants which a writer could use with a fair chance of winning appreciation and acclaim, (Cogswell 1976).

Haliburton’s *Sam Slick* (1836–1838) stories have been used by, for example, Dollinger (2010) and Chambers (2008) to establish a Loyalist connection to Nova Scotia and thus Canadian English. New (2003, 61) praises McCulloch for also being “adept at handling dialect,” though admits that Haliburton’s “handling of dialect [was] more firmly in control” than McCulloch’s. That being said, McCulloch was both Scottish and living in a part of Nova Scotian with heavy Scottish immigration. His *The Letters of Mephibosheth Stepsure* is a series of 25 letters published in the Halifax *Acadian Reader* satirizing the foibles and dialect of Nova Scotians, both Loyalist-born and Scottish-born, and employs different forms of literary dialect spelling for each of these characters.

The question is therefore, do the *McKinnon Texts* show the same incidence of the four variables under investigation as the Loyalist Haliburton or the Scottish McCulloch, and perhaps even more germane, which characters in these works employ each variant (the Loyalists or the Scots)?

### 4.2 Storyteller corpus

The second corpus of Cape Breton speech comes from recordings made of older Cape Bretoners in the 1970s and early 1980s for two different media. The first group of recordings was collected for a Canadian Broadcasting Corporation radio show called *Archie Neil’s Cape Breton* that aired for three years beginning in 1979. These recordings were of local storytellers recounting personal anecdotes or tall tales. Two hundred of these stories from 33 different storytellers were transcribed and published as *As True as I’m Sittin’ Here* (Sutcliffe and Caplan 2000). As the original recordings were not available for this project, it is from these transcribed stories that the data was extracted.

The second group of recordings were made in the late 1970s and in the 1980s by Ron Caplan, an American, who “discovered the charms of Cape Breton [in the 1970s] and believed the island’s unique character, history
and folklore should be preserved,” (Tennyson and Atwell 2002). Caplan transcribed and edited his recordings and published collections of them as *Cape Breton’s Magazine*. This magazine was published monthly, albeit irregularly, between 1972 and 1999 — though later issues reprinted interviews published in earlier volumes. The entirety of this magazine’s run was made available online in 2010. Data was also collected from several published volumes of stories from the *Cape Breton’s Magazine’s* interviews (Caplan 1980; 1988; 1991; 1996a; 1996b; 1999; 2005; 2006; 2014). These volumes include portions of the interviews that were not included in the magazine.

The transcripts published as *As True as I’m Sittin’ Here* and *Cape Breton’s Magazine* preserve idiosyncratic non-standard/local morphosyntactic, pragmatic, and lexical speech features. In fact, Caplan edited and published *As True as I’m Sittin’ Here*, so there should be similar editorial conventions in both *As True as I’m Sittin’ Here* and *Cape Breton’s Magazine*, etc. The storytellers in both *As True as I’m Sittin’ Here* and *Cape Breton’s Magazine* were identified by name and many storytellers in *As True as I’m Sittin’ Here* also told stories that appeared in *Cape Breton’s Magazine*. By searching local Cape Breton records (mostly obituaries), or from the narrative itself, it was possible to determine the (approximate) year of birth and sex of the speakers from these interviews. Despite there being many transcribed narratives from many speakers, most narratives were very short. Narratives from only 54 speakers contained one of the four variables under investigation. The 54 speakers were born between 1885 and 1943.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quotative Verbs</td>
<td>1,875</td>
</tr>
<tr>
<td>Stative Possession</td>
<td>188</td>
</tr>
<tr>
<td>Deontic Modality</td>
<td>49</td>
</tr>
<tr>
<td>Future Temporal Reference</td>
<td>558</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2,670</strong></td>
</tr>
</tbody>
</table>

### 4.3 Steelworker corpus

The third corpus of data used in this project is a set of 212 transcribed interviews from *The Steel Project* by the Beaton Institute in Sydney, Nova Scotia (Beaton 1991; 1994). This collection of oral histories was created during the modernization of the Sydney Steel Plant, which began in 1987, and was part of a larger project aimed at preserving a record of the “traditional” processes of steelmaking in the city. The interviews, conducted in 1989 and 1990, include discussions of the steelmaking process, but also the realities of being a steelworker more generally. Like the *Storyteller* corpus the interviewees are nearly exclusively blue collar, though unlike the storyteller corpus they are industrial workers rather than more rural farmers/fishermen, etc.

Most interviewees in the steelworker corpus are men, though some are women, all of whom were born between 1902 and 1960. Data extraction yielded 12,395 tokens across the four variables from 165 different speakers. Unfortunately there is no record of the transcription protocols for this data; however, many non-standard/vernacular morphosyntactic forms have been preserved — suggesting the transcripts were not excessively edited.
Table 4.3: Tokens extracted from the Steelworker corpus.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quotative Verbs</td>
<td>3,222</td>
</tr>
<tr>
<td>Stative Possession</td>
<td>3,356</td>
</tr>
<tr>
<td>Deontic Modality</td>
<td>2,224</td>
</tr>
<tr>
<td>Future Temporal Reference</td>
<td>3,593</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12,395</strong></td>
</tr>
</tbody>
</table>

4.4 Corpus of Post-Industrial Cape Breton English

The most recent data used for this study comes from the Corpus of Post-Industrial Cape Breton English (hereafter the Post-Industrial corpus, Gardner 2010b; 2013a). This corpus of 100 standard sociolinguistic interviews was collected by the author between 2009 and 2011 and has been used to explore variation and change for phonetic/phonological features (Gardner 2013a for slit-fricative (t,d), Gardner 2009; 2010a for low-back vowel fronting, Gardner 2010b for Canadian Raising; Gardner and Childs 2011; Roeder and Gardner 2013; Gardner, Roeder, and Childs 2016 for the Canadian Shift) and discourse pragmatic features (Childs and Gardner 2011 for discourse marker b’y) in Cape Breton English.

Across the four variables 3,559 tokens were extracted from 97 speakers in the Post-Industrial corpus. Together the three spoken-data corpora (Storyteller, Steelworker and Post-Industrial) will be collectively referred to as the Extended Cape Breton English Corpus.

Table 4.4: Tokens extracted from the Post Industrial corpus.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quotative Verbs</td>
<td>1,594</td>
</tr>
<tr>
<td>Stative Possession</td>
<td>1,004</td>
</tr>
<tr>
<td>Deontic Modality</td>
<td>321</td>
</tr>
<tr>
<td>Future Temporal Reference</td>
<td>640</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3,559</strong></td>
</tr>
</tbody>
</table>

The initial project for which the Post-Industrial corpus was collected in 2009, (Gardner 2010b), compared adolescent groups within a single high school in Industrial Cape Breton. For this reason there are a disproportionate number of speakers born in the late 1980s and early 1990s. Older speakers were added to the corpus through subsequent fieldwork in 2011. Data collection was guided by Labov (1984), Milroy and Gordon (2003), and Tagliamonte (2006a).
Chapter 5

Age, Sex, Social Status, etc.

The three overriding social categories considered by variationist sociolinguists when studying modern industrialized societies are age, sex or gender, and social status, usually defined as social class (Chambers 2002, 349). In this section I will outline how the three corpora of non-literary Cape Breton English presented in Chapter 4 are stratified along these three social categories, and also how these categories will be operationalized in the four case studies.

5.1 Age

Of the three social categories, age is the primary social correlate of linguistic change (Chambers 2002, 349). Excluding the McKinnon Texts, the full Extended Corpus of Cape Breton English spans 121 years of apparent time. Figures 5.1 shows the number of speakers from whose data tokens were extracted for the four case studies. Men are represented by a male symbol (♂); women are represented by a female symbol (♀). The 96 speakers from the Storyteller corpus (purple) were born between 1873 and 1940, though most were born in the early 1900s. This data was collected in the 1970s and early 1980s. The 165 speakers from the Steelworker corpus (orange) were born between 1902 and 1960. This data was collected in 1989 and 1990. Finally, the 97 speakers from the Post-Industrial corpus (green) were born between 1915 and 1999. This data was collected in 2009 and 2011. Figure 5.2 shows the number of tokens across all four case studies (a total of 17,333) by year of birth of the speaker.

Given the distribution of speakers by year of birth within each corpus and across corpora, it is evident that speaker age and corpus are not independent. Not only was the Storyteller corpus collected earlier, its speakers were born earlier, and it contains the only speakers born prior to 1902. The Post-Industrial corpus was collected most recently and includes the youngest speakers, including all speakers born after 1960. To overcome this non-independence, for each of the case studies the relative probability of a given dependent variable was tested using conditional inference recursive partitioning trees with both year of birth and corpus (also sex, see below) as test parameters. Conditional inference recursive partitioning trees (Hothorn, Hornik, and Zeileis 2006; Strobl, Malley, and Tutz 2009; Tagliamonte and Baayen 2012; Hothorn and Zeileis 2015) are especially useful for disentangling the combined effects of non-independent factors. The results of these tests revealed what I will call “shock points” within the variable system — or rather, here, certain combinations of year of birth and corpus where the likelihood of the application value of the dependent variable is significantly different between groups. It is based on these shock points that the data was partitioned, with each partition being tested independently and compared. The details of each of these tests are outlined within each case study chapter.
This blending of real time (intercorpus) and apparent time (intracorpus) data must also be kept in mind when interpreting patterns of variation and change for each of the case studies. For example, the *be like* quotative does
not appear in either the Storyteller or the Steelworker corpus, yet is attested among speakers born in the 1940s and 1950s in the Post-Industrial corpus. This is exactly the time period in which these three corpora overlap. This suggests that there has been community-wide generational change (as evidenced by the apparent time rise in the Post-Industrial corpus), but also points to individual lifespan change (as suggested by the use of *be like* by those born in the 1940s and 1950s in only the most recently collected data). For future temporal reference *be going to*, on the other hand, there is little difference across corpora among speakers in the overlapping years, but there is a difference (at least among males) in the use of the innovation among speakers born before versus after 1948. This points to generational change, but not lifespan change (or at the very least lifespan change that cannot be confirmed as statistically significantly different from chance variation).

In summary, to assess language change, the relationship between variable patterns and measures of time, which includes year of birth, corpus, and in some cases, the combination of these two parameters, was tested and where significant differences were found, partitions in the data were made. This provides the opportunity to assess whether changes in the overall frequency of a variant coincide with changes in variable patterns. It also provides the opportunity to explore at what points in time the community shows significant generational cleaves, as reflected in language.

### 5.2 Sex

Categorizing individuals into “females” and “males” has long been standard practice in the social sciences, including studying patterns of linguistic variation and change (Cheshire 2002, 243). There has been a frequent finding that women adhere to language norms more closely than men when those norms are overtly prescribed, but less closely than men when they are not. This is Labov’s (2001, Ch. 11.4) “gender paradox.” If a variant has been evaluated positively within the community or labelled “standard”, women tend to use it more frequently. This also results in women using older, conservative variants in contexts where innovations are viewed negatively. If there is a variable for which there has been no evaluation of one variant more positively than another, women tend to use the newer variant more frequently. Generally this results is men using older, often more conservative variants.\(^1\) This is an indication of the fact that “male speakers” and “female speakers” are composite groups (Eckert 2014, etc.). Indeed, Labov’s (2001, Ch. 11.4) resolution to the gender paradox (i.e., that women lead changes from above and from below) was the recognition of the diversity within the macro level category of sex: some women lead changes from below and some women lead changes from above. Many women do both.

Scrutinizing gender-based variation in Cape Breton beyond the macro-level category of sex is outside the purview of the current inquiry; moreover, an exploration of how multiple identity categories intersect among Cape Bretoners, and how identity practices relate to language is undoubtedly a worthwhile topic for future study. That being said, categorizing speakers broadly by sex is important — at least as a test parameter — because this social categorization seems to be inextricably linked to language change generally, and has also been identified as a relevant factor in past studies of quotation, stative possession, deontic modality, and future temporal reference specifically.

Figure 5.3 shows that the three corpora are not evenly stratified by sex. There are many more men in the Storyteller and Steelworker corpora than women. This is, unfortunately, an uncontrollable characteristic of the

---

\(^1\) In many vernacular speech communities and for peripheral varieties of English “traditional” variants are frequently negatively evaluated as “non-standard,” so it is often the case that an innovation and the positively evaluated form overlap, resulting in women using a higher frequency of non-local forms and men using a higher-frequency of local forms — though see Childs et al. (2010) for a contrasting scenario.
data. It may be possible in the future to add more female speakers to the Storyteller corpus as new interviews are published. The Steelworker corpus, on the other hand, is a fixed set of interviews, of which only 6 are with women. The Post-Industrial corpus, a set of sociolinguistic interviews collected by the author for this project, purposely comprises a more even number of men and women. In fact, women outnumber men in the Post-Industrial corpus. The wildly uneven distribution of male and female speakers across the three corpora results in non-independence between sex and corpus, the consequence of which is that sex and corpus could not be considered in the same regression analyses.

Male and Female Speakers by Corpus

![Graph showing distribution of male and female speakers by corpus.](image)

Figure 5.3: Speakers by sex and corpus.

![Graph showing distribution of speakers by sex, corpus and year of birth.](image)

Figure 5.4: Speakers by sex, corpus and year of birth

Figure 5.4 shows how men and women are distributed in apparent time. While men outnumber women generally, both sexes are similarly distributed across apparent time. The oldest man was born in 1878, while the
oldest woman was born in 1885. The youngest man and youngest woman were both born in 1999. So, while sex and corpus are non-independent, and corpus and year of birth are non-independent, sex and year of birth are relatively independent.

Across the four case studies sex is operationalized in two ways. First, sex was included as a parameter in the preliminary conditional inference recursive partitioning trees, which were used for determining time-measure partitions in the data. Sex was included as a parameter because of its non-independence from corpus. Second, sex was tested as a parameter within each partition, as was year of birth, because within each time measure partition these parameters are independent. The inclusion of sex and year of birth within each partition also allowed each partition to be maximally comparable to existing literature on the variables in each of the case studies.

5.3 Social Status

The third major correlate of variation and change is social status. Two of the most common indices for social status in sociolinguistic research are occupation type and level of education. Generally, speakers who work in professional or white-collar jobs (e.g., managers, doctors, lawyers, accountants, etc.), have higher social status within a community than speakers who work in manual or blue-collar jobs (e.g., steelworkers, miners, postal carriers, janitors, etc.) (Chambers 2009, 41). Likewise, speakers with some form of post-secondary education have a higher status in a community than those who have no post-secondary education (Labov 2001, 60). In variationist sociolinguistic research speakers’ social status is important to determine because in Western societies social status almost always correlates with linguistic variation (Trudgill 1974b): high social status speakers tend to use higher rates of forms considered standard and low social status speakers tend to use higher rates of forms considered non-standard.2 Once a form is overtly evaluated as being more standard/formal/proper, speakers can also choose to employ or avoid this form in order to fit in with or sound like speakers with a particular social status.

Social status is also linked to locality. The fount of standard language, generally synonymous with the speech of those in positions of political, economic, and social power and authority, is usually the seat of that power and authority: cities, not towns. The dominant bloc institutions that codify and reinforce language standards, as well as the stewards of written English (i.e., the media, publishers, etc.), which those bloc institutions name as their model (Lippi-Green 1997, 64), are all also generally located in urban not rural locales. Thus, in the context of diffusion of urban changes to more rural locales, or the convergence of peripheral varieties with the norms of the centre, Labov (2011:437) suggests that the “polarity of rural/urban” may come to replace the dimension of social class in language change. A command of standard forms becomes both a gateway to, and marker of, social status within peripheral communities (which in turn can paradoxically preserve traditional speech features among those who view this change as hegemonic). Those with high social status within a peripheral community, often also those with the means to create and maintain stronger network ties to the centre, employ a greater number of standard and therefore non-local forms (see Milroy 1980). Those with lower social status within a peripheral community employ a greater frequency of locally conservative, and thus non-standard, forms.

It is important to note that no resident of Industrial Cape Breton would consider themselves to live in a rural community, though many who live elsewhere in Canada, especially its big cities, may. Industrial Cape Breton is

---

2 Though, it is usually the case that forms are considered standard or non-standard because of the social status of those who employ them with the greatest frequency, rather than the other way around.
defined by being specifically industrial and not rural, in opposition to the rest of Cape Breton. Social hierarchies within the community over the 20th century were also based on an individual’s relationship to heavy industry. Professionals and industrial managers lived in one part of town, labourers lived in another. As Industrial Cape Breton has changed from an industrial to a service/tourism-based economy — with a significant amount of men and women travelling back and forth to Alberta or elsewhere to work in primary resource industries — social hierarchies within the community have re-organized. This means that although patterns of variation aligning with classic social class divisions may have been prevalent in the community in the past, in contemporary post-industrial Cape Breton, patterns of variation may instead reflect an individual’s alignment with a local or non-local identity.

Operationalizing social status across the three corpora of Cape Breton English is therefore a tricky endeavour. The Storyteller corpus was collected at a time when heavy industry still existed in Cape Breton, and the division between rural and industrial areas of Cape Breton was tangible. Its speakers are majority rural residents with no post-secondary education (though there are a few religious leaders, a doctor’s assistant, one miner, etc.). They also grew up in a much more linguistically heterogenous Cape Breton. All the speakers in the Steelworker corpus are blue-collar workers from a single steel plant with no post-secondary education. For these speakers the difference between blue-collar workers and white-collar managers is distinctly conspicuous — in fact, many discuss the social distinction between managers and workers. Every speaker born after 1960, all from the Post-Industrial corpus, has some form of post-secondary education (with the exception of those still in secondary school), and only a handful of speakers (all men) work in blue-collar jobs. Nearly all are either professionals or work in the service industry. Indices of social status are therefore highly inconsistent across corpora and age groups. For this reason measures of social status were not included in any analysis in this paper. Instead, an understanding of the differing social realities of each corpus will be considered when interpreting results, e.g., older speakers and speakers in the Storyteller corpus tend to be more rural; younger speakers and speakers from the Post-Industrial corpus tend to be more white-collar/educated.

5.4 Race, ethnicity, religion, etc.

Other social categories within a community can also be correlated with differences in frequency or grammatical patterning of sociolinguistic variables. While it is true that early 20th-century Industrial Cape Breton was very diverse, these other forms of social categorization have not been operationalized within the present inquiry. This is in part because the ethnic/racial and religious backgrounds of speakers from the Storyteller and Steelworker corpus are largely un-verifiable (though assumptions could be made based on these individuals’ surnames). Of note, with the exception of a handful of speakers (one African–Nova Scotian, two ethnically-Lebanese speakers, four ethnically Bajan speakers, one Mi’kmaq) all speakers in the Post-Industrial corpus and pictured in the Storyteller corpus appeared to be of mixed, white European heritage. Even though historically Industrial Cape Breton was very diverse, according to the 2011 National Household Survey only 2.5 percent of non-aboriginal Industrial Cape Breton residents were estimated to belong to a visible minority group (compared to 5.2 percent for all of Nova Scotia). The three most frequently reported ethnic origins were Scottish, Canadian, and English (Statistics Canada 2013).
Part III

Case Studies
Chapter 6

Quotatives

6.1 Introduction

The meteoric rise of *be like* as a verb for introducing direct quotations, as in (16) below, is “possibly the most vigorous and widespread change in the history of human language,” (Tagliamonte 2012b, 248). Researchers began noticing *be like* competing with existing forms like *say*, *think* and *tell* in the early 1980s (Butters 1982, 149) and in the last 35 years it has become globally ubiquitous. It has taken over the quotative system in the speech of young adults in Inland Canada (Tagliamonte and D’Arcy 2004; 2007a; Tagliamonte and Denis 2014; Gardner et al. 2013) and the United States (Barbieri 2007), Australia (Rodríguez-Louro 2013), New Zealand (Buchstaller and D’Arcy 2009; D’Arcy 2012), and parts of the United Kingdom (Buchstaller 2011; Cheshire et al. 2011; Durham et al. 2012; Gardner et al. 2013).

(16)  *be like*  I remember talking to a girl from Sydney Academy and she was like, “It’s weird to wear sneakers to school.”
       (Female born 1991, PI)

       *be like*  Up in Ontario they’d be like, “Sneakers? They’re running shoes!” (Female born 1961, PI)

Tagliamonte, D’Arcy, and Rodríguez-Louro (2016, 842) have called the simultaneous rise of *be like* across multiple varieties of English a “linguistic Black Swan event”—something that could not have been predicted by socio/historical linguistic theory. As part of their proposal, Tagliamonte, D’Arcy, and Rodríguez-Louro (2016) show that *be like* started to be used by speakers born in the same years in Toronto and Victoria, Canada, Perth, Australia, and Christchurch, New Zealand, suggesting a lightning-fast diffusion that could only have been possible due to rapid concomitant changes in youth culture, international travel, and communication technologies. Not only did *be like* emerge at the same time in these four communities, the internal linguistic factors constraining the use of *be like* in all four cities are identical. This is a linguistic development that is inconsistent with models of linguistic diffusion such as the wave model (Bailey et al. 1993; Labov 2003; Trudgill 1974a; 1986), the gravity model (Bailey et al. 1993; Callary 1975; Hernández Campoy 2000; Kerswill 2003; Labov 2003; Larmouth 1981; Trudgill 1974a), the cascade model (Callary 1975; Hernández Campoy 2003; Kerswill 2003; Labov 2003; Trudgill 1986), and the cultural hearth model (Horvath and Horvath 1997; 2001; 2002).

Tagliamonte, D’Arcy, and Rodríguez-Louro’s (2016) findings aside, in some communities the rise of *be like* does show the classic consequences of linguistic diffusion: simplification of complex grammatical or social constraints—at least in its initial stages within a community. These include smaller and peripheral communities like Springville, in east central Texas (Cukor-Avila 2002) and St. John’s, Newfoundland (D’Arcy 2004). This is

---

1. Butters (1982, 149) describes *be like* as a “narrative use of to be (usually followed by like).”
exactly what has been found for Cape Breton English. In this chapter an analysis of *be like* will be presented that strongly points to linguistic diffusion. The adoption of *be like* in Cape Breton is so clearly diffusion, the present case study will serve as a benchmark in assessing whether the changes examined in the three subsequent case studies are changes brought about by transmission or diffusion.

### 6.2 Synchronic perspective

The largest and most comprehensive analysis of quotative verbs worldwide is Gardner et al. (2013, 525 speakers, N=15,827), which combined data from past studies of quotative verbs in York, U.K. and Toronto, Canada, (Tagliamonte and Hudson 1999; Tagliamonte and D’Arcy 2004; Durham et al. 2012) with data previously unexamined with respect to quotative verbs and with new data collected in the two cities in 2013. The conclusion of the study was that *be like* had reached the saturation point in Toronto English. Among the youngest speakers in the study (those younger than 30 interviewed in 2013), the frequency of *be like* versus other quotative verbs was about 70 percent, and for some speakers within this group it was well over 90 percent or categorical. The linguistic and social constraints on *be like* use, observed among the middle-aged speakers in the study and reported elsewhere for *be like*, were completely or nearly completely levelled among the young speakers. The oldest speakers in the study (those older than 60 interviewed between 2002 and 2004) used no *be like* at all, and had rates of *say* of about 80 percent, indicating a complete overhaul of the Toronto quotative system in less than 70 years.

Tagliamonte and Denis (2014) report on the spread of *be like* to three communities outside of Toronto: Lakefield, Burnt River, and Belleville, Ontario (see Map 5 on xvii). Contra Tagliamonte, D’Arcy, and Rodríguez-Louro’s (2016) evidence of simultaneous adoption and identical constraints across communities, Tagliamonte and Denis (2014) find the classic consequences of diffusion, including temporal lag and the simplification of linguistic constraints (relative to Toronto). The authors propose that *be like* was not diffused to these communities as a variant with a full set of linguistic constraints; rather, the fixed expression *I’m like* diffused to the communities, and from there the *be like* quotative expanded within the quotative system in a process akin to grammaticalization. Elsewhere in Canada, D’Arcy (2004) reports that the 14 adolescent female speakers in her *St. John’s Youth English Corpus* (collected in 1999–2000) use the *be like* quotative 62 percent of the time (N=184). Henley et al.’s (2008) analysis of 12 men and 12 women aged 19–25 recorded eight years later in St. John’s showed *be like* had increased in real-time in the community to 73 percent (N=1,169). This growth is similar to the change in rates of use of *be like* by <30 year-old Toronto speakers recorded in 2002-2004 and <30 year-old Toronto speakers recorded in 2012/2013 presented by Gardner et al. (2013). The linguistic constraints governing the use of the *be like* variant in St. John’s are also very similar to Toronto. While it is usually the case that innovations like quotative *be like* begin in focal areas like Toronto and then diffuse to more peripheral communities, like those in Southern Ontario and St. John’s, the similarity in linguistic conditioning in St. John’s conversely suggests linguistic drift. Dion and Poplack (2005) find that among similarly linguistically isolated Québec anglophones there are not only strikingly similar rates of *be like* use to Ontario speakers, but identical internal linguistic conditioning for the use of the variant, again suggesting drift. Much like Cape Breton English, Québec English and St. John’s English have been classified as autonomous varieties of Canadian English (Boberg 2010). Both Cape Breton English and St. John’s English have also been labelled as quickly — though idiosyncratically — converging with Inland Canadian English (for example, Clarke 2010; Keifte and Kay-Raining Bird 2010). The present study will determine if, like speakers in Québec and St. John’s, Cape Breton speakers are using quotative verbs just like their counterparts in Toronto, suggesting drift, or have a simplified system compared to Toronto speakers — just like other Southern Ontario speakers — suggesting imperfect replication.
through diffusion.

6.3 Methods

The methods for examining quotative verbs are intentionally modelled after both Gardner et al. (2016) and Tagliamonte and D’Arcy (2004). As in these studies, all instances of verbs introducing direct quotations were extracted from the Storyteller, Steelworker, and Post-Industrial corpora. The full suite of verbs found across the corpora are listed, with examples, in (17). These verbs include semantically neutral options like say and tell, options that emphasize the quality of the quoted speech like holler, scream and yell, options that highlight the illocutionary force of the quoted speech like vow, options that indicate the quoted speech is performative like go or do, options that position the quoted speech within a broader discourse routine like ask, answer, reply, and hear, and finally, options that indicate that the quoted speech is internal or self-reflected like decide, realize, feel, think, and wonder.

(17) answer The cow answered, “We had a beautiful summer. We were eight of us in a huge, big pasture.” (Male born 1908, ST)
ask We came back and I go back to the gate and I asked the fellows there, “Where’s the Whitney Pier?” (Male born 1910, ST)
be And in the wintertime, boy, that thing would be, “Sssssssss!”; the steam coming out of it. (Man born 1921, PI)
be like And the owner was like, “Ah, tell him wait till Tuesday and we’ll cut the roof out, put a sunroof into it, and drive it to Cape Breton for him.” (Male born 1983, PI)
call You could literally hear my dad calling a “Lindsey!” (Female born 1980, PI)
decide And for some reason or other we decided, “Naw, we may as well stay here.” (Female born 1946, PI)
do You’re talking business and things will come up and then you do, you know, “Oh gosh, yeah, there was a building there and it’s gone now.” (Female born 1958, PI)
explain I explained to Gordie, “Oh, his wrist is sore.” (Male born 1985, PI)
feel And that hurt me more, I felt, “Oh gee! They knew very well I couldn’t do that.” (Female born 1916, ST)
figure He figures “Okay, well, I’m going to stomp these guys, because they’re just posers.” Right? (Male born 1989, PI)
go The smell wasn’t a smell that you went, “Ew, that’s an awful smell!” (Male born 1955, SW)
hear And all of a sudden I hear, “Where you be hailing from, blah blah blah?” (Female born 1956, PI)
holler He was hollering, “For Jesus sake, get over and give me some room!” (Male born 1907, ST)
realize And you realize, “Oh yeah, where am I?” (Female born 1961, PI)
reply I replied, “Oh really?” (Male born 1964PI)
say I always say it. And then after I say it, I say, “Oh my god I can’t believe I just said that [laughs]. ” (Female born 1991, PI)
scream We were standing there screaming, “Get them out, get them out!” (Male born 1908, ST)
tell That’s what I told him, I told him I’m not going to borrow ten bucks for you.” (Male born 1952, SW)
think I just think, “Ah shit, I could die if I don’t get unburied soon.” (Male born 1985, PI)
write He wrote, “Don’t come, because there’s nothing here.” (Male born 1948, SW)
yell We yelled, “Surprise!” (Female born 1991, PI)

Verbs for introducing direct quotations, like say and think, can also introduce indirect quotations, as in (18), or some other clause, as in (19). These instances of quotative verbs were not extracted or included in the analysis.

(18) And he says that he makes over a hundred thousand dollars a year. And he has, like, a newer truck, but it’s not even brand new. (Male born 1983, PI)

(19) Well my brother, actually, he works at Scotsburn and he — I think I might end up working there this summer, he said he could probably get me a job ten bucks an hour, so that would be awesome. (Male born 1992, PI)
Instances of *be like* with existential *it* as a subject, like (20), were extracted but were not included in the analysis of *be like*'s internal constraints because other quotative verbs, like *say* and *think*, do not occur with an existential subject. Tokens where *it* represents an inanimate object acting as the subject of a quotative verb would have been included in the analysis, but no construction like this occurred in the data.

(20) It’s like, “Oh, it’s Saturday? Okay. Well, let’s get loaded then go to the bar.” (Male born 1989, PI)

Finally, instances of direct quotation with no overt quotative verb were also extracted. These types of constructions often occur when a speaker is reporting the details of a conversation, as in (21):

(21) They said “Yeah. Yeah, but why would you talk to them?” \(\lor\) “They’re friends of mine I’m driving.” \(\lor\) “Really? Well, where are you from?” \(\lor\) “The Pier.” \(\lor\) “Oh my God. I never would have thought you were from the Pier.” Well, I had to be held back at that point. (Female born 1967, PI)

### 6.4 Results

#### 6.4.1 Trajectories of Change

##### 6.4.1.1 All Quotatives

A total of 6691 quotative contexts were extracted from the *Storyteller*, *Steelworker* and *Post-Industrial* corpora. Figure 6.1 clearly demonstrates that *say* (the dashed green line with triangles) is the majority variant for speakers born before the 1960s.

![Distribution of quotative verbs in Cape Breton English](image)

Figure 6.1: Distribution of quotative verbs in Cape Breton English — variants by decade of birth. (Speakers born 1878–1899 collapsed.)

The lower rates of *say* use among the oldest speakers (those born before 1910) and those born in the 1960s and 1970s is not due to the increased rates of another quotative verb variant, but instead a greater use of the zero quotative (dot-dashed pink line with asterisks). Zero quotatives are used most often in contexts where speakers are recounting a conversation and among the data from speakers born in the 1900s, 1960s and 1970s there is a greater number of tokens that occur while speakers are recounting conversations, as in (21) above. The 1960s is the first birth decade in which 100 percent of speakers are part of the *Post-Industrial* corpus. As described in Chapter 4, this corpus is comprised of sociolinguistic interviews. Part of the interview schedule included...
asking speakers to recount arguments or disagreements they experienced. Therefore, this corpus includes many examples of speakers quoting lengthy dyadic exchanges which often occurred with a zero quotative after first reference. Aside from say, the zero quotative, and be like (solid blue line with squares) among those born in the 1980s and 1990s, no other quotative verb is used more than 8 percent of the time by speakers in each decade. Table 6.1 shows the distribution of the variants across all corpora.

Table 6.1: Quotation in Cape Breton English – variants by decade of birth.

<table>
<thead>
<tr>
<th></th>
<th>be like</th>
<th>say</th>
<th>go</th>
<th>think</th>
<th>tell</th>
<th>ask</th>
<th>other</th>
<th>zero</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1878-1899</td>
<td>0</td>
<td>137</td>
<td>74%</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>4%</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>1900s</td>
<td>0</td>
<td>712</td>
<td>79%</td>
<td>0</td>
<td>7</td>
<td>1%</td>
<td>54</td>
<td>6%</td>
<td>37</td>
</tr>
<tr>
<td>1910s</td>
<td>0</td>
<td>774</td>
<td>88%</td>
<td>1</td>
<td>0%</td>
<td>8</td>
<td>1%</td>
<td>16</td>
<td>2%</td>
</tr>
<tr>
<td>1920s</td>
<td>0</td>
<td>1641</td>
<td>92%</td>
<td>8</td>
<td>0%</td>
<td>2</td>
<td>0%</td>
<td>21</td>
<td>1%</td>
</tr>
<tr>
<td>1930s</td>
<td>0</td>
<td>449</td>
<td>95%</td>
<td>6</td>
<td>1%</td>
<td>1</td>
<td>0%</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>1940s</td>
<td>10</td>
<td>901</td>
<td>89%</td>
<td>3</td>
<td>0%</td>
<td>8</td>
<td>1%</td>
<td>5</td>
<td>0%</td>
</tr>
<tr>
<td>1950s</td>
<td>11</td>
<td>441</td>
<td>85%</td>
<td>21</td>
<td>4%</td>
<td>3</td>
<td>1%</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>1960s</td>
<td>17</td>
<td>99</td>
<td>48%</td>
<td>7</td>
<td>3%</td>
<td>17</td>
<td>8%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1970s</td>
<td>1</td>
<td>31</td>
<td>8%</td>
<td>0</td>
<td>1</td>
<td>8%</td>
<td>0</td>
<td>0</td>
<td>1%</td>
</tr>
<tr>
<td>1980s</td>
<td>220</td>
<td>57</td>
<td>15%</td>
<td>13</td>
<td>4%</td>
<td>4</td>
<td>1%</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>1990s</td>
<td>198</td>
<td>57</td>
<td>17%</td>
<td>9</td>
<td>3%</td>
<td>2</td>
<td>1%</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>457</td>
<td>5275</td>
<td>68</td>
<td>53</td>
<td>109</td>
<td>79</td>
<td>91</td>
<td>559</td>
<td>6691</td>
</tr>
</tbody>
</table>

Marginal quotative variants include verbs like think, tell, ask and go. The “other” category includes verbs like call, holler, swear, and reply, none which are used more than 1 percent of the time by any age group.

6.4.1.2 Go

The early literature on be like found that the variant was part of a system in which another innovative variant, go, appeared to be growing rapidly in use. For example, Tagliamonte and Hudson (1999, 58) found that among a group of university students in Ottawa interviewed in 1995 go was used 22 percent of the time, compared to 13 percent for be like and 36 percent for say. In the same study, university students from York, U.K., interviewed in 1996 used go 18 percent of the time — the same frequency with which they used be like — compared to using say 31 percent of the time. In Cape Breton go is a very minor player in the quotative system for all speakers. Of the 6,691 tokens extracted from the data, there were only 68 tokens of go. The go quotative appears earlier in both apparent and real time in the Cape Breton data compared to be like; and, while it starts increasing from 1 or less precent usage to 3–4 percent usage among those born in the 1950s and 1960s — be like makes the same jump in frequency among people born at the same time — go remains stable at 3–4 percent usage in subsequent decades of apparent time. Be like, on the other hand, skyrockets. Figure 6.2 shows the rates of be like (solid blue line with squares) and go (dashed orange line with circles) by speaker decade of birth, from the 1880s to 1990s, and also by corpus: Storyteller; Steelworker; and Post-Industrial.

D’Arcy’s (2012) analysis of young females in St. John’s interviewed in 1999 and 2000 and Henley et al.’s (2008) analysis of twenty-something men and women in St. John’s interviewed in 2008 both show incredibly low rates of go. It is likely that the use of go as a versatile quotative for introducing direct speech (rather than gestures or non-lexical sounds) either did not diffuse, or simply did not take off on Canada’s East Coast. In fact, of the 68 go tokens, one third (N = 22) introduce non-speech sounds, like the six tokens from 1930, in (22), of one man telling a ghost story; non-lexical speech sounds, like (23); or gestures, like (24). And of all the tokens
that introduce these types of quotations ($N = 33$), all are either \textit{go} ($N = 22$), \textit{be like} ($N = 8$), \textit{hear} ($N = 1$) or zero ($N = 2$). This suggests that, unlike for Ottawa students in 1995, \textit{go} is primarily a quotative specialized for reporting non-speech sounds, i.e., auditory impressions, non-lexical exclamations, or gestures.

(22)

Non-speech sound  The last fellow going through the door saw the mat left lying on the hearth \textit{go}, “\textit{Swoooosh},” up the chimney! (Male born 1930, ST)

Non-speech sound  All of a sudden they could hear the sewing machine upstairs \textit{go}, “\textit{Clickity-clack, clickity-clack}!” (Male born 1930, ST)

Non-speech sound  They thought that very strange and they went upstairs and the sewing machine had indeed been \textit{going}, “\textit{Clickity-clack, clickity-clack}!” (Male born 1930)

Non-speech sound  The chest \textit{went}, “\textit{Woosh},” sliding along underground. (Male born 1930, ST)

Non-speech sound  If you were to speak, then the chest would \textit{go}, “\textit{Chooooooo},” sliding along underground and you wouldn’t find it that time. (Male born 1930, ST)

Non-speech sound  Just as he was lying down to take his \textit{first} rest on his own land on the Island of Cape Breton, he could see coming over the rise of the hill a large black dog with fiery eyes and the dog \textit{went}, “\textit{Gruff! Gruff! Gruff}!” (Male born 1930, ST).

(23)

Non-lexical sound  The smell wasn’t a smell that you \textit{went}, “\textit{Ew}, that’s an awful smell.” (Male born 1955, SW)

Non-lexical sound  St. Peter would read a bit and then, every once in a while \textit{go}, “\textit{Mmmmm, mmmm}.” (Male born 1910, ST)

(24)  Gesture  Here’s what the superintendent does when he comes in to look at a job: [stands up and pantomimes]. He \textit{goes}, [\textit{looks and nods}]. He’s trying to convince you that he knows what’s going on. (Male born 1940, SW)

Gesture  So I walked into the pen, and I took her bridle off, and then I walked out of the pen and I closed the gate and then she comes up on the gate and she \textit{went}, [\textit{makes sad face}], and she just closed her eyes and she \textit{went}, “Why are you leaving me?” (Female born 1999, PI)

Quotative \textit{go} does appear in Gray (2007) as a local speech feature. The quotations introduced by \textit{go} in (25) are all direct speech — an environment among Cape Bretoners to which \textit{go} appears not to have grammaticalized.
This suggests that, while it may not occur frequently in Cape Breton speech, quotative go is likely part of the
suite of non-standard features believed by residents to be part of Cape Breton speech (see Section 3.2). This
further suggests that Cape Bretoners are aware of quotative systems external to the island and are therefore not
isolated from possible diffusion.

(25) GOES — A multi-faceted word used to connect long sentences when gossipin to your nay burr. “Well I goes, she’s such
a snot, and she goes, yah well yer a wussey, and I goes, well I ain’t go got a bun in the oven and she goes”…And so on and so on. (Gray 2007, 24)

6.4.1.3 Be like

Be like’s jump from negligibility in the system to being used about 60 percent of the time strongly suggests
diffusion. The sudden rise in apparent time of be like among speakers born in the 1980s and 1990s points to an
“off-the-shelf” (Milroy 2007) adoption of the be like quotative, rather than a local development. Transmission
of be like from older generations to younger generations is impossible given that older generations do not use
be like. The oldest speakers who use be like are three females born in 1946 but recorded in 2011. Despite a
substantial number of tokens from speakers born earlier, there is no evidence of be like used as a quotative. For
this reason drift also seems unlikely, especially also given the rapidity of be like’s usurping of say as the dominant
quotative variant. Diffusion is further supported by the lack of be like tokens from speakers born in the 1940s,
1950s, and 1960s from the Storyteller and Steelworker corpora, which were collected 20 and 30 years before the
Post-Industrial corpus (see Figure 6.2 and Table 6.2). If be like were an independent parallel development in Cape
Breton English, or an inherited low-level feature, we would expect to see evidence of be like in the older, larger
datasets too — not just the new data. This points to be like’s entry into Cape Breton sometime after those born
in the 1940s, 1950s, and 1960s were adults (e.g., 1980), and indicates that speakers born in these decades are
adult-adopters of be like rather than speakers who acquired the feature natively from their caregivers.

Table 6.2: Quotation in Cape Breton English — be like by decade of birth and corpus.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>% be like</td>
<td>Total N</td>
<td>% be like</td>
<td>Total N</td>
</tr>
<tr>
<td>1940s</td>
<td>4.7% 228</td>
<td>0% 779</td>
<td>0% 8</td>
</tr>
<tr>
<td>1950s</td>
<td>6.6% 168</td>
<td>0% 351</td>
<td>0</td>
</tr>
<tr>
<td>1960s</td>
<td>9.0% 188</td>
<td>0% 17</td>
<td>0</td>
</tr>
<tr>
<td>1970s</td>
<td>7.6% 13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>6.5% 597</td>
<td>0% 1147</td>
<td>0% 8</td>
</tr>
</tbody>
</table>

In the be like literature, three main conditioning factors for be like use are consistently attested: content of
the quotation, grammatical person of the subject, and temporal reference of the quotative verb (Ferrara and Bell
1995; Tagliamonte and Hudson 1999; Tagliamonte and D’Arcy 2004; Tagliamonte and Denis 2014; Tagliamonte,
D’Arcy, and Rodriguez-Louro 2016, etc.). The first of these three conditioning factors concerns the type of
information being introduced by the quotative verb. Be like is believed to have entered the quotative system as a
way to describe hypothetical speech, like internal thought as in (26), and then extended to introduce all types of
dialogue, including direct speech, as in (27), and writing, as in (28), (Romaine and Lange 1991, 262; Tagliamonte
2012b, 248).

(26)
Internal thought I remember walking over it and wondering "What used to be here?" (Male born 1945, PI)
Internal thought I said to myself, "I won’t see anything because I’m not doing this on my own." (Male born 1910, SW)
Internal thought And I was trying to swim and I was like "What’s going on here? I can’t move. Like, I’m not moving." (Male born 1983, PI)

(27)

Direct speech Oh yeah, there’s the pony [points to photograph]. I used to ride on her back. I used to go for the cows with her. I’d say, "Bobby, we’ll go for the cows." (Female born 1916, PI)
Direct speech They gave me a map, and they’re like, "Alright, so you’re going to turn left at the Bulk Barn advertisement." (Male born 1991, PI)

(28)

Writing There was a sign there that said "Clearance," whatever it was. (Male born 1964, PI)
Writing I can’t wait to tell him that. Like you have no idea. I just want to text him right now and be like, “Ah-ha! You were the problem child, not me!” (Female born 1987, PI)
Writing No, they all just kind of like staring, and she texted me, and she’s like, “This is so awkward.” (Female born 1991, PI)

Figure 6.3: Distribution of be like by content of the quotation in Cape Breton English —Post-Industrial corpus only.

Figure 6.3 shows the distribution of be like tokens by content and and year of birth condensed into two groups: those born in the 1940s to 1970s and those born in the 1980s and 1990s. This data comes from the Post-Industrial corpus only. As expected, for both the older and younger age groups there is a clear preference for be like in internal thought contexts.

The second common pattern observed for be like use is the contrast between first person and third person grammatical subjects (Tagliamonte and D’Arcy 2007a, 209; Tagliamonte 2012b, 249). Be like has been found to be used at higher frequencies with first person subjects, as in (29), relative to third person subjects as in (30).

(29) 1st person And I called mom later that day and I was like, "Oh, I was at this zoo." (Female born 1993, PI)
1st person We drove down there, and we were like, "What the hell is going on here?" (Female born 1987, PI)

(30) 3rd person Kent comes around and flashes a badge at us, says "Mall security. Time to leave." (Male born 1991, PI)
3rd person You had your Grade 4 teacher saying "No, you can’t say that. Can’t say that. That’s not proper English." (Male born 1989, PI)
There are two proposed explanations for why first person favours be like relative to third person within the be like literature. The first explanation is that be like begins as a way to describe internal thought (Ferrara and Bell 1995, 283; etc. Tagliamonte and D’Arcy 2007a, 212) and in narration the narrator is more likely to be aware of and express his or her own internal thoughts — so this is the environment where be like begins to grow first, and then it spreads to other contexts through grammatical extension (one of the mechanisms of grammaticalization, see Ferrara and Bell 1995; Romaine and Lange 1991). Alternatively, it has been argued that first person subjects were an initial favouring environment for be like for the reasons above, but that be like has grown in use with all subject types at a constant rate (Gardner et al. 2016). The second explanation for be like being used more frequently with first person subjects is one put forward by Tagliamonte and Denis (2014). The authors hypothesize that when be like diffused to three communities outside of Toronto it did so as a fixed construction, I’m like, and it is from this construction that the first generation of be like users inferred a variable rule for be like’s use, with first-person subjects and present-tense contexts favouring be like. These two proposals do not conflict; the former describes a dialect-internal actuation of be like, while the latter is an explanation for communities where be like diffused.

Figure 6.4 shows the relative frequency of be like use by grammatical person for the two age groups that use be like. As noted previously, tokens of be like with an existential it subject have been excluded from these frequencies because this context is not variable. For each age group the rate of be like use with first person and third person subjects is quite similar. There is less of a difference in frequency between first and third person subject tokens than between either and other subject tokens (i.e., with second person or no/null subjects) for 1980s-1990s speakers, or between internal thought and direct speech tokens for all speakers. The weak contrast between first and third person subjects differentiates the Cape Breton quotative system from the current Toronto quotative system, and from quotative systems observed in other English-speaking countries. It does however align with D’Arcy’s (2004) findings from St. John’s young females, for whom there is not a significant difference in be like use between first person subject and third person subject contexts — though Henley et al. (2008) do find a significant person effect among their male and female speakers born in the same birth years but recorded eight years later.

The third common pattern observed for be like is that the form is initially favoured with present tense morphology when the verb also has a past-temporal reference (Blyth, Recktenwald, and Wang 1990; Singler 2001). This combination of tense and temporal reference is called “narrative” or “historic” present (Tagliamonte and D’Arcy 2007a, 204) and is often employed when narrating past events. For example, the actions of speaking described by the verb says in both examples of (31) clearly happened in the past, despite their present tense morphology. In the first example, the past-tense morphology attached to the verb told in the coordinate clause, as well as the broader context of the narrative, indicate this. Likewise, in the second example the speaker clearly equates the temporal reference of says with the past temporal reference of said.

(31)

| Historical present | I remember one night distinctly, my dad, um, after dinner, he told the girls to, he says, “You girls clear the table. And you boys stay right there.” (Male born 1958, PI) |
| Historical present | The first thing he said to me, he says, “I’m taking the buzz.” (Male born 1955, PI) |

Most quotative verbs with present tense morphology extracted from the three corpora were used in this way, mostly because in the sociolinguistic interviews that make up the Post-Industrial corpus speakers were asked to recount past events, and for the Storyteller and Steelworker corpora speakers were asked to recount stories
or discuss the history of the steel plant. Some of these past events and anecdotes are recounted using the past tense, though other tenses, including present tense with a present temporal reference or habitual meaning, are also used. Some examples are presented as (32).

(32) Simple past “Oh,” she said, “I thought that was something you were buying for the cows or the horses.” (Male born 1918, PI)

Simple present, I always say it. And then after I say it, I say, “Oh my god I can’t believe I just said that.” (Female born 1991, PI)

(habitual reference)

Future, And he will say, “That, that right there. What do you call that?” (Female born 1946, PI)

(with will)

Modal, In that area, yeah. That, you know, and you might drive by there one night, and go, “Oh, that’s so-and-so’s car there.” (Female born 1958, PI)

(Past habitual, I mean, the engineer would be in with you and he’d tell you, “Now, brake it easy. There’s no need - easy so it will slow down.” (Male born 1947, SW)

Figure 6.4 shows the distribution of be like in four contexts: with simple past tense morphology and past temporal reference (Simple Past), with present tense morphology and past-temporal reference or a habitual meaning (Simple Present), and finally, all other contexts (e.g., future temporal reference, modals, infinitives, etc.). As with the other figures above, be like with existential subjects was not included in this distribution because the context is not variable. Figure 6.5 shows a clear four-way distinction among speakers born in the 1980s and 1990s, with a higher rate of be like in historical present contexts, compared to past tense, other present tense, and other contexts. Almost all of the be like tokens from speakers born before 1980 are in historical present contexts.

In their Toronto and York study Gardner et al. (2013) find the quotative verb tense to be the strongest constraint for every speaker group, and though all internal linguistic constraints were found to be levelling among young Toronto speakers, the tense constraint was the least levelled. In Toronto the authors found the three way tense/morphology distinction of historical present, >past >present

but in York the authors found only a temporal reference distinction, with past-temporal reference (whether with present tense or past-tense morphology)

2. Most studies of be like exclude "other" contexts.

Figure 6.4: Distribution of be like by grammatical person of the subject in Cape Breton English — Post-Industrial corpus only.
favouring be like, and present-temporal reference disfavouring be like. In other words, there was not a significant difference between simple past and historical present, but there was a significant difference between both and present. This fact was part of the evidence Gardner et al. (2013) used to argue that York had a less complex variable quotative system than Toronto.3 Though, Tagliamonte and Denis (2014) found that in two of the outlying Southern Ontario communities, despite there being no content-of-the-quotation constraint, the distinction between past and present was still strong — however, this constraint interacted highly with the person constraint and was in essence a consequence of I’m like being a fixed construction that diffused to the community. Among the young St. John’s females studied by D’Arcy (2004) be like was used 100 percent of the time in historical present contexts. There was also a significant difference between present temporal reference contexts, which favoured be like, and simple past contexts. This suggests that for St. John’s the relevant distinction is perhaps morphological rather than based on temporal reference. D’Arcy (2004) collected her data in 1999-2000. Eight years later, by the time of Henley et al.’s (2008) study, the tense constraint had levelled among St. John’s speakers. Between the three systems — Toronto, York and St. John’s — Cape Breton most closely resembles Toronto. This is unexpected given that Cape Breton aligns with St. John’s and not Toronto with respect to the grammatical person constraint.

6.4.2 Linguistic Conditioning

Table 6.4 (p. 64) presents two mixed-effects logistic regression analyses of the probability of be like use, one for speakers born in the 1940s–1970s and one for speakers born in the 1980s–1990s, with tense/temporal reference, subject grammatical person, content of the quotation and speaker sex included as test parameters. The analyses use the lme4 package available in that statistical modelling environment R (Bates, Maechler, et al. 2015; R Core

3. Though a reevaluation of the same data by the same authors, Gardner et al. (2016), testing specifically for evidence of the Constant Rate Effect (Kroch 1989; 1994), lead the authors to posit that be like arose via drift in both communities.
Team (2015). As before, existential *it* subjects have been excluded. Table 6.4 also accounts for individual speakers’ differing propensities for using *be like* and differing number of tokens within the data by including speaker as a potential random effect.\(^4\) The log-odds reported in the *lme4* model have been converted into probability factor weights using the inverse logit function, as factor weights are the conventional method of expressing likelihoods/probability in language variation and change literature.\(^5\) Factor weights range between 0 and 1. Factor weights greater than .5 favour the application value (here *be like*) relative to the mean, those below .5 disfavour the application value relative to the mean. Factors weights of exactly .5 are equal to the mean, and therefore neither favour nor disfavour the application value. Square brackets around factor weights indicate non-significant contrasts. The *lme4* package (Bates, Maechler, et al. 2015) tests the difference between individual levels of a parameter and the overall mean probability by using sum contrast coding. Within the variationist literature the widely-used variable rule program *Goldvarb* (Sankoff, Tagliamonte, and Smith 2015), as well as the *lme4*-based *Rbrul* script (Johnson 2009), perform a similar/the same analysis.

A Wald \(\chi^2\) test (Fox and Weisberg 2011) was used to determine which factor groups or parameters in the model were significant and which were not (Table 6.3).\(^6\) The \(p\)-values for each parameter in Table 6.3 indicate the probability that the magnitude of effect of a parameter (represented by the \(\chi^2\) statistic) is equal to zero given each parameter’s degrees of freedom. Parameters where \(p < 0.05\) can be considered to have a magnitude of effect on the variation that is statistically significantly different from null. The corollary is that there is a significant contrast between the levels of that parameter (i.e., the factors of that factor group) and the overall mean.

Table 6.3: Analysis of deviance — Wald \(\chi^2\) test for full model. *be like* for quotation in Cape Breton English.

<table>
<thead>
<tr>
<th>Parameter (factor group)</th>
<th>df</th>
<th>1940s–1970s (\chi^2)</th>
<th>(p)-value</th>
<th>1980s–1990s (\chi^2)</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tense/Temporal Reference</td>
<td>2/1</td>
<td>26.15 (2.1 \times 10^{-06})</td>
<td>***</td>
<td>79.47 (2.2 \times 10^{-16})</td>
<td>***</td>
</tr>
<tr>
<td>Subject Grammatical Person</td>
<td>2</td>
<td>1.64 (4.4 \times 10^{-01})</td>
<td>***</td>
<td>38.33 (4.8 \times 10^{-09})</td>
<td>***</td>
</tr>
<tr>
<td>Content of the Quotation</td>
<td>1</td>
<td>0.12 (7.2 \times 10^{-01})</td>
<td>***</td>
<td>7.81 (5.2 \times 10^{-03})</td>
<td>**</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>3.95 (4.6 \times 10^{-02})</td>
<td>*</td>
<td>0.34 (5.6 \times 10^{-01})</td>
<td>*</td>
</tr>
</tbody>
</table>

\(^{*} p < 0.001, ^{**} p < 0.01, ^{*} p < 0.05\)

The \(\chi^2\) value is the test statistic. Degrees of freedom (df) is the number levels for a given parameter minus 1. The \(p\)-value is determined by comparing the test statistic and the df to the \(\chi^2\) distribution.\(^1\) There were no simple present tokens among 1940s–1970s speakers, so the Tense/Temporal Reference parameter has differing degrees of freedom between age cohorts.

For factor groups with three or four factors, like tense/temporal reference and grammatical person of the subject, a secondary mixed effects multiple regression model was built to test the difference between individual levels rather than between those levels and the overall mean. The secondary models used here go one step fur-

---

\(^4\) See Johnson (2009) and Tagliamonte and Baayen (2012) for the benefit of considering both fixed and random effects in sociolinguistic analyses.


\(^6\) The Wald \(\chi^2\) test iteratively adds and removes each parameter of the model and compares how well each iteration fits the distribution of the data. If a parameter is found to be significant, it is interpreted as adding explanatory value. If a parameter is not significant, its contribution is superfluous to the understanding of the data and can be set aside. In this way, the Wald \(\chi^2\) test is very similar to the step-up, step-up down procedure implemented by *Goldvarb* (Paoliillo 2002; Tagliamonte 2006a). The result of the Wald \(\chi^2\) test reveals what combination of original parameters make the most statistically parsimonious model, or rather, which combination of factors best accounts for the variation in the data.
ther by testing the difference between individual levels of a parameter themselves using treatment or “dummy” coding. This type of analysis is more common within the broader scientific literature. For parameters with only two levels both sum contrast and treatment coding lead to identical conclusions — levels that are significantly different from each other will also be significantly different from the mean probability (which is the mean of the means of each parameter). For parameters with more than two levels, however, treatment coding can reveal additional information, e.g., whether a two-way, three-way, four-way, etc., contrast exists within a parameter.  

The first model presented in Table 6.4 shows that for the older speakers, born in the 1940s–1970s, the only significant predictors of be like use are tense/temporal reference of the quotative verb and sex of the speaker. Be like is favoured with historical present and by women. The significant contrast for tense/temporal reference is between historical present and simple past/other tokens. There were no be like tokens from this group with simple present morphology and a non-past temporal reference.

The second model in Table 6.4 shows that for the younger speakers, born after 1980, tense/temporal reference and content of the quotation are both statistically significant. For tense-temporal reference there is a clear four-way contrast between historical present, simple past, simple present, and other tense/temporal reference constructions. This pattern aligns with tense/temporal reference patterning in other communities in which be like has been studied. Likewise the contrast between internal thought and other kinds of quotations is a significant one. The model does indicate that the subject grammatical person factor group is a significant predictor of the variation; however, within this factor group the significant distinction is between first and third person subjects, which both favour be like, and other subjects, which disfavour be like. In other studies which find grammatical person of the subject to be significant, the contrast that is significant is first person versus third person. Other contexts are not included in those models. This must be taken into account when comparing this model to past findings. Finally, among the younger speakers there is no sex effect.

### 6.5 Discussion

The above results are an attestation of the rapid rise of be like in yet one more variety of English — here one on the northeastern periphery of the North American macro-dialect area. Toronto speakers younger than 30 recorded in 2012/2013 by Gardner et al. (2013) use be like at a rate of about 68 percent. Cape Breton speakers born in the same time period and recorded in 2009–2011 use be like 56 percent of the time. If all the “other” contexts (second person subjects, other tenses, etc.) are put aside (as they are in the literature), the rate among Toronto speakers is 87 percent and among Cape Breton speakers is about 77 percent. While this 10–12 percent difference might seem large, they are relatively proportional when the rates of the previous generation are taken into account. In Toronto, speakers born in the 1940–1970s, recorded in 2002–2004, use be like at a rate of 17 percent while the speakers born in same decades in Cape Breton only use be like 4 percent of the time. So while Cape Breton may lag in the overall rate of be like use relative to Toronto, the actual explosion of use of be like in Cape Breton is more rapid in apparent time.

The difference in rates between speakers born in the 1940s–1970s in Cape Breton and Toronto is the first piece of evidence that be like diffused to Cape Breton rather than it being an independent development. The lack of be like in the Storyteller and Steelworker corpora for speakers born in these same decades also strongly suggests that the use of be like among these speakers is due to adoption of the form later in life. One caveat, however, is that the Steelworker and Storyteller corpora are almost entirely working-class men who are (according to

---

7. See Tagliamonte (2012b), Tagliamonte, Denis, and Gardner (2014), and Gardner (2016) for further discussion of the usefulness of combining the two approaches.

8. Just looking at the 1960s and 1970s, the rate in Toronto is 26 percent compared to Cape Breton’s 7 percent.
Table 6.4: Mixed effects logistic regression — contribution of factors to the probability of *be like* for quotation in the *Post-Industrial* corpus (c. 2009–2011).

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>1940s–1970s</th>
<th>1980s–1990s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Be like</strong> in Cape Breton English</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>aic</strong></td>
<td>171</td>
<td>529</td>
</tr>
<tr>
<td><strong>input</strong></td>
<td>.0091</td>
<td>.59</td>
</tr>
<tr>
<td><strong>total N</strong></td>
<td>579</td>
<td>683</td>
</tr>
</tbody>
</table>

**TENSE/TEMPORAL REFERENCE**

<table>
<thead>
<tr>
<th></th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical Present</td>
<td>.92</td>
<td>33</td>
<td>33</td>
<td>.90</td>
<td>93</td>
<td>199</td>
</tr>
<tr>
<td>Simple Past</td>
<td>.30</td>
<td>3</td>
<td>213</td>
<td>.56</td>
<td>69</td>
<td>172</td>
</tr>
<tr>
<td>Present</td>
<td>KO</td>
<td>0</td>
<td>31</td>
<td>.30</td>
<td>49</td>
<td>88</td>
</tr>
<tr>
<td>Other</td>
<td>.17</td>
<td>1</td>
<td>302</td>
<td>.17</td>
<td>15</td>
<td>224</td>
</tr>
<tr>
<td><strong>range</strong></td>
<td>75</td>
<td></td>
<td></td>
<td>73</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SUBJECT GRAMMATICAL PERSON**

<table>
<thead>
<tr>
<th></th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Person</td>
<td>[.62]</td>
<td>7</td>
<td>168</td>
<td>.69</td>
<td>75</td>
<td>236</td>
</tr>
<tr>
<td>3rd Person</td>
<td>[.55]</td>
<td>3</td>
<td>209</td>
<td>.67</td>
<td>68</td>
<td>283</td>
</tr>
<tr>
<td>Other</td>
<td>[.33]</td>
<td>1</td>
<td>202</td>
<td>.18</td>
<td>7</td>
<td>164</td>
</tr>
<tr>
<td><strong>range</strong></td>
<td></td>
<td></td>
<td></td>
<td>51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CONTENT OF THE QUOTATION**

<table>
<thead>
<tr>
<th></th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Thought</td>
<td>[.53]</td>
<td>11</td>
<td>56</td>
<td>.69</td>
<td>65</td>
<td>83</td>
</tr>
<tr>
<td>Direct Speech/Other</td>
<td>[.47]</td>
<td>3</td>
<td>523</td>
<td>.35</td>
<td>52</td>
<td>600</td>
</tr>
<tr>
<td><strong>range</strong></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SEX**

<table>
<thead>
<tr>
<th></th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>.75</td>
<td>5</td>
<td>379</td>
<td>[.54]</td>
<td>58</td>
<td>314</td>
</tr>
<tr>
<td>Male</td>
<td>.25</td>
<td>1</td>
<td>200</td>
<td>[.46]</td>
<td>55</td>
<td>369</td>
</tr>
<tr>
<td><strong>range</strong></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Random Effects**

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
<th>N</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>2.45</td>
<td>29</td>
<td>1.94</td>
<td>39</td>
</tr>
</tbody>
</table>

Sum contrast coding.

Aic, input, and factor weights for significant factor groups derived from most parsimonious model. Factor weights for non-significant factor groups derived from full model.

Factor weights converted from log-odds. $p = \frac{e^x}{(1+e^x)}$.

Factor weights with non-significant contrasts indicated with [ ].

Correlation of fixed effects

1940s–1970s $r < |0.271|$; 1980s–1990s $r < |0.344|$.

Additional treatment contrast coding used to test orthogonal contrasts for parameters (factor groups) with more than two levels (factors) using most parsimonious model.

Significant contrasts indicated with horizontal lines.
the sociolinguistic literature) the least likely to use innovative features (though see Chapter 9), while the *Post-Industrial* corpus includes professional and working-class male and female speakers, and is thus expected to have at least some innovative features.

In much of the past literature on *be like* women have shown significantly higher rates of *be like* compared to men (Tagliamonte 2012b, 249). This, of course, is consistent with Labov’s (2001, 266, and elsewhere) observation that women lead in the use of incoming variants. But if the presence of women within the dataset is what results in there being *be like* within the dataset, there should be a significant difference between the rates of *be like* use between men and women. There is not.

While there is a significant difference in *be like* use between men and women born between 1940–1970 (5 percent for women, 1 percent for men) in Table 6.4, this difference disappears if the *it’s like* tokens are included in the calculations. Normally *it’s like* is not included when calculating rates and modelling *be like* use because only *be like* can occur with existential subjects, and also because it is a categorical third person context. However, because *it’s like* can replace *he says* or *I say*, as in (34) on p. 68, the difference between when speakers choose any kind of *be like* construction vs. some other kind of quoting strategy can be evaluated. Looking at all direct quoting strategies, i.e., including the *it’s like* tokens, the rate of *be like* use for women born in 1940s–1970s is 7 percent and for men born in the same decades is 6 percent, a contrast found to be non-significant (*p* = 0.36). In fact, with *it’s like* included the difference between male and female speakers is non-significant for speakers born in each decade from 1940s to 1990s (though, there is no data from 1970s males).

![Distribution of *be like* by sex in Cape Breton English](image)

**Figure 6.6: Distribution of *be like* by sex in Cape Breton English – *Post-Industrial* corpus**

Tagliamonte (2012, 250, based on Tagliamonte and Hudson 1999, 159; Ferrara and Bell 1995) presents a set of predictions for the grammaticalization of *be like*. She argues (again, based on Ferrara and Bell 1995) that in *be like’s* initial states, females use *be like* more than males and it occurs at higher frequencies with first person subjects and internal dialogue. If *be like* has expanded to third person and direct speech contexts at a later stage of its development, then sex differentiation also ought to be neutralized. In the Cape Breton data *be like* is found in third person contexts and with direct speech, so maybe the lack of a strong sex effect is to be expected. This suggests that *be like* arrived in Cape Breton at a later stage of its development. This strongly points to diffusion of *be like* to Cape Breton from a source where it had already reached a later stage of grammaticalization and young Cape Bretoners replicated the constraints associated with this later stage.
Table 6.5: Distribution of *it’s like* across corpora — adapted and expanded from D’Arcy (2004, 334).

<table>
<thead>
<tr>
<th>Location</th>
<th>N quotatives</th>
<th>% <em>it’s like</em></th>
<th>% <em>be like</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cukor-Avila (2002)</td>
<td>1,346</td>
<td>0.2</td>
<td>13</td>
</tr>
<tr>
<td>Tagliamonte and Hudson (1999)</td>
<td>612</td>
<td>0.2</td>
<td>13</td>
</tr>
<tr>
<td>Tagliamonte and Hudson (1999)</td>
<td>665</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Ferrara and Bell (1995)</td>
<td>417</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>Tagliamonte and D’Arcy (2004)</td>
<td>2,058</td>
<td>5</td>
<td>58</td>
</tr>
<tr>
<td>Gardner et al. (2013)*</td>
<td>1,970</td>
<td>6</td>
<td>58</td>
</tr>
<tr>
<td>D’Arcy (2004)</td>
<td>184</td>
<td>8</td>
<td>62</td>
</tr>
<tr>
<td>Gardner et al. (2013)*</td>
<td>1,939</td>
<td>2</td>
<td>69</td>
</tr>
<tr>
<td>Henley et al. (2008)</td>
<td>1,169</td>
<td>4</td>
<td>73</td>
</tr>
<tr>
<td>Cape Breton 1940s–1970s</td>
<td>597</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Cape Breton 1980s–1990s</td>
<td>718</td>
<td>5</td>
<td>58</td>
</tr>
</tbody>
</table>

* = only data collected in 2012/2013

D’Arcy (2004) makes a similar claim for St. John’s. She argues that the *be like* quotative is more “deeply embedded” in the speech of St. John’s young females in 1999 and 2000 than had been seen previously. Part of her argumentation is that St. John’s shows higher rates of *it’s like*,9 which she claims (based on Singler 2001, 260) represents a later stage of grammaticalization of *be like* within the quotative system; however, the lower rates of *it’s like* reported by Tagliamonte and D’Arcy (2004) and Gardner et al. (2013), and between D’Arcy (2004) and Henley et al. (2008) complicates this interpretation because, as they feature speakers born and/or recorded after D’Arcy’s (2004) informants, they would be expected to have 8 percent or higher *it’s like* use. Both post-1980s and pre-1980s Cape Bretoners show a rate of *it’s like* comparable to the rate reported for Toronto by Tagliamonte and D’Arcy (2004) and York by Gardner et al. (2013). The Toronto speakers in the corpus used by Tagliamonte and D’Arcy (2004) were recorded in 2002-2004 and ranged from teenagers to speakers in their 90s. The York speakers in the corpus used by Gardner et al. (2013) cited in Table 6.5 were recorded in 2013 and were all under 30. Gardner et al. (2013) argued, based on a number of factors, that the 2013 York speakers were about a generation behind Toronto speakers with respect to the use and grammatical patterning of *be like*, and that the quotative system in York showed signs of diffusion. It appears that Cape Breton is very similar to York in that respect, and the similar rates of *it’s like* use in both cities provides more evidence for this conclusion.

In this scenario *be like* diffused to Cape Breton sometime in the 1970s or early 1980s, a decade later than its arrival/development in Toronto. The Cape Breton speakers from the 1980s and 1990s are the first generation of native users of *be like*, all older speakers are adult adopters. The lack of a person or sex effect among the 1980s and 1990s speakers is either due to the simplification of complex constraints associated with adult-to-adult diffusion, or due to *be like* being at a later stage of development in in the source variety, in which some constraints (like the person and sex constraints) had previouslylevelled. Both explanations for these simpler constraints in Cape Breton are congruous with *be like* diffusing to Cape Breton and unlikely if *be like* were a parallel independent development (c.f., Gardner et al. 2016).

An alternative interpretation of the data is also possible. Tagliamonte and Denis (2014) use the strong interaction between tense/temporal reference and subject person to argue that the avenue through which *be like*

---

9. Here *it’s like* represents existential *it* with any tense/temporal reference: e.g., *it’s like, it was like, it’ll be like, it’d be like*, etc.
diffused to outlying Ontario communities was through the fixed expression I'm like, and that the next generation of speakers inferred a variable rule from this fixed expression. This resulted in the next generation’s favouring of be like in present-tense, first-person contexts. This variable rule developed and persists because these contexts are favoured for be like in other, more prominent, global varieties of English — the most important of which being nearby Toronto English. In Cape Breton, and perhaps in St. John’s, a similar phenomenon may have occurred.

**Distribution of be like by grammatical person of the subject, tense/temporal reference, and content of the quotation in Cape Breton English**

![Graph showing distribution of be like by grammatical person of the subject and tense/temporal reference in Cape Breton English — Post-Industrial corpus only.](image)

Figure 6.7 shows the distribution of be like by grammatical person of the subject and tense/temporal reference of the quotative verb in the Cape Breton data. There is a clear early preference for be like in historical present contexts with first person subjects, though the use of first and third person subjects to report direct speech still occurs among the oldest speakers. Among the younger speakers in all but one combination of tense/temporal reference and content type first person subjects favour be like relative to third person subjects. The one context

---

10. This aligns with Trudgill’s (2014) view that quotative be like is a form of lexical replacement rather than grammatical change, and therefore spreads like lexical items. I suggest that the spreading of be like may occur as a fixed expression, but that it is subsequently integrated into the quotative system and is subject to grammatical conditioning in ways that expressions like groovy, cowabunga or Whassup? are not.
that appears to resist *be like* is first person subjects with a present tense verb used to report direct speech. Looking closely at this context in the data, nearly all the tokens are speakers describing how they pronounce something or whether or not they use an expression, as in (33). These types of tokens are present in the data because the Post-Industrial corpus interview protocol included questions about language use. Perhaps these types of expressions will prove to be *say*’s “terminal refugium” as *be like* saturates the English quotative system.

(33) a. Right? And I can’t take that. I can’t say “talk.” (Male born 1989, PI)
c. I don’t know it’s just the way the- the ‘OO’ sound comes out it-I say, “stupid.” (Female born 1990, PI)
d. Like, I say, ’T-N’ instead of “tonight” or ‘td” instead of “today.” (Female born 1990, PI)
e. Oh my god, yes. I always say, um, “B-R-B.” I always say it. And then after I say it, I say “Oh my god I can’t believe I just said that [laughs].” (Female born 1990, PI)

The data in Figure 6.7 strongly points to *be like* being diffused to Cape Breton as the fixed expression *I’m like* just like the outlying Southern Ontario communities examined by Tagliamonte and Denis (2014). But if this is the case, what accounts for the overall lack of significant effect between first and third person subjects in the statistical analysis? If *be like* diffused to Cape Breton as the lexical construction *I’m like*, should not there be a subsequent favouring of *be like* in first person contexts? The answer is twofold. First, direct speech makes up the bulk of the data (90 percent of tokens among the older speakers, 88 percent among the younger speakers), and non-historical present tense constructions are relatively infrequent (0.5 percent of tokens among older speakers, 13 percent among the younger speakers). Even though the patterns observed in internal thought or present tense contexts may be extreme with respect to a person effect, they do not outweigh (in the statistical analysis) the overall pattern found in the more plentiful contexts — that first and third person subjects have relatively similar rates of *be like*, even if first person has higher rates of the variant.

The second part of the answer lies in Figure 6.8, which shows the distribution of *be like* for each decade of speakers who use it, but with *it’s like* constructions included in the distributions. The orange bars represent the proportion of the *be like* tokens that are *it’s like*, the green portion represents *I’m like*, and the blue portion represents all other *be like* constructions. What is clear is that the absolute percentage of *it’s like* tokens remains fairly stable across age groups, but the relative percentage of *it’s like* tokens is larger among those born in the 1940s and 1950s than among those born later. Among some speakers, like the male born in 1958 in (34), *it’s like* is used nearly categorically.

(34) *it’s like* You know, everybody else would be drifting off. It was like “Where you all going?” It’s like “Home for supper.” It’s like “Oh yeah, okay. Me too.” So, and then, so there was no real boundary. And I mean, you could go quite far, and you’d be up in Ashby at ten years of age, visiting a buddy. But somebody knew you were there sort of thing, right? I’d tell my brother Alan, you know, it’s like “Tell Mom I’m going over Johnny’s house.” And “Yeah, okay. Whatever.” And, it seemed like nobody really worried about you. And then, of course, we all get curfews when we were fourteen, fifteen years old. It’s like, “You be home by twelve o’clock” or whatever, right? (Male born 1958, PI)

I therefore argue that the initial avenue of diffusion of *be like* to Cape Breton was not just *I’m like*, as it was in outlying Southern Ontario communities, but instead *I’m like* and *it’s like*. The variable rule that was inferred by subsequent generations as children was such that there was no distinction between first-person (from *I*) and third-person (from *it*) subjects, but that both favoured *be like* relative to other kinds of subjects. The other significant constraints — the favouring of historical present and internal thought contexts — would also have been inferred from both the tense of *it’s like* and *I’m like* (present) but also the usage of the two: narrating past events in which the internal psychological experience of the narrator was an integral part of how the story is told (cf. Buchstaller and D’Arcy 2009).

Figure 6.8: Distribution of be like including it’s like and I’m like in Cape Breton English — expressions by decade of birth.

In this alternative scenario speakers born in the 1940s–1950s are the initial adopters of the lexical expressions I’m like and it’s like, which they acquired in the mid-1960s to late 1970s — the period when be like was being rapidly diffused worldwide (Tagliamonte, D’Arcy, and Rodriguez-Louro 2016). Despite hewing closely to off-island norms with respect to changes with long time depths (as will be shown in subsequent chapters), steel workers born during this time period — who discuss working in the steel plant immediately following high school in the Steelworker corpus — may have been less likely to come into contact with speakers beyond the Canso Causeway, as the speakers born in the 1940s and 1950s in the Post-Industrial corpus were (all the be like users in this group discuss traveling off island after high school). Unlike the cities examined by Tagliamonte, D’Arcy, and Rodriguez-Louro (2016), which had a /fluid movement of young speakers in the 1960s and 1970s, Cape Breton since 1960 has been characterized by depopulation. Only some young people, like those born in the 1940s and 1950s in the Post-Industrial data, left and then returned. These speakers are the agents of change within the community — at least for this innovation. These agents of change were young Cape Breton speakers interacting with their off-island peers in the late 1960s and 1970s as university students or travellers. These agents of change brought the lexical expressions I’m like and it’s like back to the island with them. The continued use of these expressions was spurred by contact with off-island tourists, and with former Cape Breton residents (cousins, friends, etc.) now living in other speech communities as a consequence of deindustrialization and depopulation. The next generation of speakers — the 1970s/1980s children of those born in the 1940s and 1950s — who have also had contact with off-island cousins and tourists and hear be like from popular media, inferred a set of grammatical rules for using be like as a fully grammaticalized quotative expression. This more restricted pathway of diffusion in Cape Breton resulted in a delay of about a decade before be like use skyrocketed within the community. The constraints for be like use that were inferred by the first generation of be like users in Cape Breton are nearly identical to those for be like in other, more prominent, global varieties of English.
among speakers born in the same birth years.\textsuperscript{12} By situating the diffusion of \textit{be like} in the mid-1960s to late 1970s, rather than the late 1970s and 1980s, this scenario for the diffusion of \textit{be like} to Cape Breton includes the community within the global emergence of \textit{be like}. Furthermore, by categorizing the diffusion as lexical, this scenario accounts for the unique pathways of diffusion of a depopulating, peripheral community; it aligns Cape Breton with other Canadian communities (like Lakefield, Burnt River and Belleville) outside the country’s major metropolitan areas; and, it accounts for the decade delay in \textit{be like}'s rise in Cape Breton. Like the first scenario, described above, this second scenario is incongruous with \textit{be like} being a parallel independent development in Cape Breton (c.f., Gardner et al. 2016).

\section*{6.6 Conclusions}

The evidence presented in the previous section indicates \textit{be like} likely diffused to Cape Breton. For \textit{be like} use to grow from null among those born before the 1940s, to 7 percent among those born between 1940–1970, to about 58 percent among those born after 1980 (and over 80 percent in some contexts), indicates that young speakers could not have received \textit{be like} through an unbroken chain of adult to child transmission. That not one single token of the 4,922 quotative tokens extracted the \textit{Storyteller} and \textit{Steelworker} corpora and from the \textit{Post-Industrial} from speakers born before 1940 was \textit{be like} suggests the form simply did not exist earlier in the community. \textit{Be like} is undeniably a change from above. In other peripheral Canadian communities (e.g., anglophone Québec) the internal constraints for \textit{be like} use match those of Toronto, suggesting linguistic drift. In Cape Breton this is not the case. The internal constraints for \textit{be like} at its actuation in the community do not match the earliest speakers in Toronto, they are simpler. This fact points to either of two scenarios. Either \textit{be like} diffused to Cape Breton from a source that was at a late stage of \textit{be like} development, where \textit{be like} had already begun to be widely used in non-first-person and non-internal dialogue contexts, and where sex differentiation had been neutralized. Or, \textit{be like} diffused to Cape Breton as two fixed expressions: \textit{I'm like} and \textit{it's like}, and from these fixed expressions linguistic constraints were inferred by subsequent generations. As these expressions were used equally by male and female initial adopters, a sex effect did not develop — despite the general tendency for females to use a higher frequency of innovative forms relative to males of the same age (Labov 2001, 262, and elsewhere). Evidence for both of these scenarios includes the lack of a significant subject constraint for \textit{be like} use among all \textit{be like} users, as well as nearly identically rates of \textit{be like} use among men and women.

\textit{Be like} is an important point of comparison when examining the changes in progress in the ensuing three case studies. \textit{Be like} is a prototypical diffused change. A specific group of speakers can be identified as the first generation of native speakers and there is a clear loss of grammatical complexity in the constraints that govern the variation between \textit{be like} and other forms, relative to the ostensible source of the diffusion: Inland Canada. If the rise of certain variants of the grammatical variables examined in the next three chapters — e.g., \textit{have} for stative possession, \textit{have to} for deontic modality, and \textit{be going to} for future temporal reference — show the same hallmark signs of diffusion, it may be concluded that, like \textit{be like}, these variants’ current use in Cape Breton is likely due to geographic diffusion. If, on the other hand, the analysis of these variants do not show a simplification of complex grammatical constraints, relative to Inland Canada or Scotland, their use in Cape Breton is likely due to either transmission or drift.

From a broader perspective, this chapter represents the largest analysis of quotative verbs in a Canadian speech community outside of Toronto. It adds to the wide body of literature which finds tense/temporal refer-

\textsuperscript{12} In Toronto English, for example, \textit{be like} has saturated the quotative system, so constraints like subject person are levelling (Gardner et al. 2013). For example, the rate of \textit{be like} use with first and third person subjects or between men and women is becoming smaller as the rate of use of \textit{be like} in these contexts or among these speakers approaches 100 percent.
ence and quotation content to strongly condition variation between *be like* and other forms. It also conforms to the observation that in communities in which *be like* begins to be robustly used later diachronically — like St. John’s and Cape Breton — the elsewhere strong contrast between rates of use in first person and third person contexts and between men and women are not significant. The present analysis also bolsters the proposal put forth by Tagliamonte and Denis (2014) that in some communities *be like* was diffused as a lexical expression and that in subsequent generations grammatical constraints were inferred. While this is not the only scenario the data supports, it is perhaps a superior interpretation because it additionally accounts for the unique pathways of diffusion to a peripheral, depopulating community while simultaneous allowing for Cape Breton’s participation in global trends as they occur. Unlike the rapid simultaneous adoption of *be like* among young speakers in the 1960s and 1970s without loss of grammatical complexity, observed in larger centres by Tagliamonte, D’Arcy, and Rodriguez-Louro (2016), the spread of *be like* to Cape Breton does not subvert the conventional explanations for the diffusion of linguistic innovations. It instead can easily conform to the wave, gravity, cascade and cultural hearth models of linguistic change. This suggests, as do Tagliamonte and Denis’s (2014) findings, that while *be like*’s meteoric rise worldwide may defy expectations in some speech communities, in others it follows predictable patterns of sociolinguistic diffusion.
Chapter 7

Stative Possessives

7.1 Introduction

Claims that distinctive features of Cape Breton English originate exclusively from its Scots or Scottish Gaelic settlers may have some validity with respect to phonology, as discussed earlier (Chapter 3.2). In the following three case studies the validity of this claim with respect to morphosyntactic features will be tested.

To begin to answer this question this chapter examines variation among stative possessive verbs in Cape Breton English. Variation between *have*, *have got* and *got* (as in (35) below) is especially germane to investigating the origins of Cape Breton English because the trajectory of change for this variable is not the same in Canada as in the British Isles. Furthermore, *have got/got* had not arrived in northern Ireland or Scotland by the time Cape Breton’s Celtic settlers emigrated (as described in Chapter 3.1.1). Therefore, *have* vs. *have got* vs. *got* is a perfect test site for determining if the community grammar in Cape Breton has its origin in Scottish/Irish English and British English, or in Loyalist English, like Inland Canada.

7.1.1 The Variable

In Cape Breton English, like other varieties of English, stative possession is usually expressed using one of three forms: *have*, *have got*, and *got*, as in (35).\(^1\)

(35)  
*have* I have three sisters. (Male born 1991, PI)  
*have got* He has got a shaved head. (Female born 1991, PI)  
*got* Because sometime, well, if you got the beginning of cataracts, or anything like that, sometimes it makes a big difference, you know. (Male born 1942, PI)

The phonetically-reduced forms of the verb *have* (’s/’ve) are also included in this variation, as in (36).

(36)  
*have got* It’s got two grey cats in it and dustballs the size of rats. (Female born 1946, PI)  
*have got* Now today you’d make freezer jam. I’ve got a little bit of it home. (Female born 1922, PI)  
*have* I’ve friends from Mexico. (Female born 1991, PI)

It is not difficult to find variation between these three forms in Cape Breton. For example, (37) shows that the variants occur even within the same utterances of individual speakers — here, from males born in 1991, 1952, and 1926:

---

1. Stative possession can also be expressed with the verb *possess*, *hold* (a degree), *be endowed with*, etc. These either did not occur or occurred at exceedingly low levels (one or two times) in the data and are not discussed in this chapter.
a. It’s two stories, I think. It’s about 1,300 square feet, I think. And we got about two-and-a-half acres of backyard — half of which my dad actually mows. And we have a pool table in the basement. (Male born 1991, PI)
b. She has her Master’s in library science. And she is some kind of an adjudicator, or a head librarian at the, ah, I think they call it Capital Centre, in Halifax now. I think that’s pretty close to a description. She got three or four letters behind her name and that. But, I’m, I haven’t got that. (Male born 1952, PI)
c. I had a good friend of mine was an engineer. He was the first engineer on the caster. He was only young. We used to bowl together and play darts together — a group of us, you see. I got to know this fellow pretty well. I used to go fishing with him. Anyway, I talked to him, with him different times what was going on. He was a good guy, no doubt. He had spent a year in Austria, very close to Hitler’s birth place, where they bought that caster. He used to tell me, “The system over there is a lot superior to the one here.” When he first said it to me, I said, “What do you mean? They have a better way of working?” “Oh, yes,” he said, “Over there all the workers got bicycles, the foremen got motorcycles, and the officials have cars.” [laughs] Well I tell you, we used to be at it hammer and tongs. (Male born 1926, SW)

Variation also occurs in printed material. The have and have got forms appear in Photograph 4 and 5 — two local advertisements for Better Health Centre and Woolco, which were both found several times in the same publications as the oral history narratives that comprise the Storyteller corpus.

Photograph 4: Better Health Centre Advertisement, Cape Breton’s Magazine, March 1, 1973, p. 15

Photograph 5: Woolco Advertisement, Cape Breton’s Magazine, March 1, 1973, p. 2

The got variant also appears in the two-volume local folk dictionary Da Mudder Tongue and Anudder Mudder as an example of a Cape Breton idiom or vocabulary item, and in the example sentences of other entries:

(38) YAGOTENNY? — You have to say this phrase with no pauses between words. It means do you have any. (Gray 2006, 60)
The local form is *You got any?* and the translation of this form is *do you have any?* This suggests that the bare *got* possessive is evaluated as both non-standard and local (these two ideas tend to be confounded in the community, see 5.3), while the *have* possessive is evaluated as standard and/or non-local — even though, as shown below, it is the majority form in the community.

### 7.2 Historical Perspective

The oldest of the three stative possessive forms is *have*, which has encoded the stative possessive meaning since the late 10th century. Starting in the 16th century *have got*, and then just *got* in the mid 19th century, gradually began replacing *have* for the expression of stative possession in British English. This pattern continues in British and non-North American Englishes (e.g., Kroch 1989; Noble 1985; Quinn 2004; Tagliamonte 2003; Tagliamonte, D’Arcy, and Jankowski 2010), where *have got* is often found as the majority form or a form on the rise (Biber et al. 1999, 466; Tagliamonte 2006b; Kroch 1989; Noble 1985; though also see Gramley and Pätzold 1992, 350). In North American English, with which the “bare” *got* form has been traditionally associated (Jespersen 1961a, 53; Quirk et al. 1985, 131–132), the conservative *have* is now on the rise in apparent time (e.g., Jankowski 2005). This is an unusual instance where a change in progress has reversed (Tagliamonte 2015, and elsewhere).

Tagliamonte and Denis (2014, 89) use Kroch’s (1989, 209) findings (based on Noble 1985) from British plays to conclude that the inland United Empire Loyalists who settled in Southern Ontario had rates of *have got* somewhere between 7 percent and 24 percent. It it likely that the mixed coastal United Empire Loyalists who settled in Cape Breton would have had a similar rate. Tagliamonte (2013a, 149), on the other hand, finds that among the older speakers she examined from Northern Ireland and northern Scotland, *have got* had yet to make significant inroads. The Scottish and Irish speakers who settled in Cape Breton would have emigrated from these homelands before *have got* had taken root there, so it would not have been possible for them to carry this feature with them to Cape Breton. Furthermore, as discussed previously, most of the Scottish settlers in Cape Breton did not speak English, so a specifically Scottish English stative possessive system being transplanted to the island is highly improbable.

The distinction between Loyalists and Scottish usage is reflected in the writing of Thomas McCulloch. McCulloch’s Alexander Scantocreesh, a Scot, never employs *have got*, while Mephibosheth Stepsure and the character of Loopy, both of Loyalist/British descent, do frequently, as in (39).

(39) a Mr. Catchem *has got* a large house; and he has already been going a good deal about, expressly for the purpose of inquiring after those who are uncomfortable at home. (McCulloch 1821–1822/1990, Letter 6 (as written by Mephibosheth Stepsure))

b Loopy, stopping a little for the purpose of offering me a great bargain of a horse, gave me a very discouraging view of the farming life. “Meph,” says he, “I’ll tell you what it is, you have got a world of hard work before you. Upon my word, the farmer has got a laborious life of it. I do assure you, it takes a great deal of toiling to maintain a family by a farm; and after all, it won’t do.” (McCulloch 1821–1822/1990, Letter 10 (as written by Mephibosheth Stepsure))

c You and I hae great matter o’thankfu’ness, that we hae been spared to be what we are. (McCulloch 1821–1822/1990, Letter 20 (as written by Alexander Scantocreesh))

d Were ye to hear the fok in our toon speakin’, or to see them at the kirk wi’ their shaises and ither bravery, ye wad think them lords and leddies: but the maist o’ them haena a bawbee in their pouch; and at hame, some hae scarcely a caff bed to sleep on. (McCulloch 1821–1822/1990, Letter 20 (as written by Alexander Scantocreesh))

Haliburton’s Yankee character Sam Slick and Slick’s Bluenoser interlocutors also employs innovative *have got* frequently, as well as the older *have*, as can be seen in the following examples from just one sketch of the 33 published.
7.2.1 The McKinnon Texts

Given the above historical and literary perspective, it is somewhat strange that McKinnon’s (1852) Loyalist protagonists never use have got/got, while the “colourful characters,” who show other Scottish/Irish features or who are otherwise considered low-status (the social position of most Scottish/Irish Cape Bretoners in 1850s) do employ have got/got. For example, the following from the Backlander in The Midnight Murders:

(42) Al haven’t got knowledge for that, dus toime, you please. (McKinnon 1852, 4-5)

Or from a criminal sailor in the same novel:

(43) “Well,” cried one of the men [a sailor], “he can only tell you the same story — we’ve got nothing to do with it. He was going to be put on board the — and we were paid for it — it’s not of our affair — and we can’t help it if he gets away for if you rescue him — we done our share, and you won’t let us do the rest. Come along, boys, we’ve got the money at any rate.” (McKinnon 1852, 88)

McKinnon, a Loyalist, was clearly aware of the incoming have got/got variant and likely made the common prescriptivist false assumption that innovative variants are “improper” or “non-standard” and are thus only used by the lower strata of society, and also used universally by all speakers who occupy this social position. It is also possible that in the 1850s got forms were already beginning to be thought of unfavourably — a prescription that begins to take hold in the last half of the 19th century in North America (as summarized by Rice 1932). If McKinnon was sensitive to this stigmatization, it is no surprise that he uses have got/got to signal that his colourful characters are of low-status and different from his 100 percent have-using protagonists. This is consistent with Krug’s (2000, 61–62) analysis of have got/got in Dickens’ Oliver Twist (1837/38), where the have got/got forms, according to Krug, are part of a suite of non-standard phonological and grammatical features used by Dickens to characterize non-standard speech. McKinnon’s (1852) mistake of putting have got/got in the mouths of Scot-
tish/Irish speakers, if anything, showcases the stratification of colonial Cape Breton, with high-status Loyalists like McKinnon having little direct interaction with low-status coasters or backlanders.²

### 7.3 Synchronic perspective

Tagliamonte (2006b, 317) reports that in contemporary Toronto English have is the predominant form (77 percent), especially among the youngest speakers, with both have got (18 percent) and stigmatized got (6 percent) playing minor roles in the system — especially relative to American English where got occurs 40 percent of the time in written dialogue (Jankowski 2005; Tagliamonte, D’Arcy, and Jankowski 2010).

Walker and Hoffman (2016) show even more have use. In their dataset of Toronto speakers of differing ethnic backgrounds, the authors find that overall have was used 86 percent of the time, have got only 6 percent of the time, and got 8 percent of the time. There was a significant difference in have versus have got/got use between speakers of various ethnic origins. Speakers with an Italian ethnic background and speakers with a British/Irish ethnic background were significantly more likely to use either of the have got/got forms relative to speakers with other ethnic backgrounds. Among the speakers with a British/Irish background, speakers older than 40 (c. 2010) used stative possessive have about 75 percent of the time, while speakers 18–30 (c. 2010) used stative possessive have 85 percent of the time.

In communities outside of Toronto (see Map 5 on p. xvii), Tagliamonte and Denis (2014) find that have is also the majority form. For speakers in Burnt River, Lakeview and Bellville, Ontario, have represents about 75 percent of the data, with have got representing just under 20 percent and got less than 10 percent.

In York, U.K. on the other hand Tagliamonte (2006b) finds that use of have is declining among young speakers, while have got is on the rise (see Figure 7.1). Among the peripheral northern British, northern Irish, and Scottish speakers examined by Tagliamonte (2013a), the very oldest speakers use the most have and there is a shift in

---

² Though it is also possible that rural Scottish/Irish speakers in 18th century Cape Breton may employ more have got/got forms than Sydney Loyalists, having learned the forms from visiting Yankee tourists or from relatives who had travelled back and forth to “the Boston States” in the late 19th century. This interpretation of McKinnon’s dialect writing choices still aligns with the form having Yankee/Loyalist origins.
apparent time towards have got (Tagliamonte 2013a, 149). The rates of have and have got are also aligned from south to north and from urban to most peripheral, with have favoured in communities that are more north or more peripheral. This includes the northern Irish and Scottish communities where have got/got is scarcely used among old speakers — suggesting that innovative have got began in the southern cities and is diffusing to northern towns. At present there exists neither a descriptive nor quantitative account of have vs. have got/got for Highland or Hebridean Scottish English — the dialect areas that 19th Century Cape Breton immigrants arrived from. Tagliamonte’s (2013a) apparent-time analysis of stative possessives in a small corpus collected in Buckie, in northeast Scotland, is the best available proxy for Highland Scottish English. Her analysis of stative possessives among older speakers in Cullybackey and Portavogie, Northern Ireland is also the only analysis available describing the variable use of stative possessives in Irish English.

In New Zealand English, another transported variety with initial immigrant populations (British, Scottish and Irish) similar to Cape Breton, Quinn (2004) finds that among speakers born in the mid- to late-1800s have got was used just under 40 percent of the time in positive present-tense possessive contexts. And as it is in British English (but not North American English), have got is on the rise in New Zealand. Among those born between 1900 and 1978, Quinn (2004) finds have got is the majority form, occurring between 60 percent and 90 percent of the time. Quinn links the rise in have got to immigrants from the south of England, where have got is attested earlier than Antipodean immigration, or from Australia (whose ancestors were from south-eastern England).

### 7.4 Methods

The aim of this dissertation is to assess whether Cape Breton English is "more or less identical with the English of the Scottish Highlands" (Trudgill 2004, 7) using stative possessives, deontic modality verbs, and future temporal reference expressions as a test cases. Assessing genetic relationships between varieties by comparing variable systems is referred to as "comparative sociolinguistics" (as described in Section 2). The method for extracting, coding and analyzing tokens from the Cape Breton data for stative possessives was therefore modelled on the methods used by Tagliamonte (2013a) for Scotland/Ireland and Tagliamonte and Denis (2014), etc. for Southern

3. Buckie English derives from Scots, while Highland and Hebridean Scottish English do not.

4. The Irish-Newfoundlanders who emigrated to Cape Breton would have spoken a version of southern Irish English. While not an ideal proxy, the data from Cullybackey and Portavogie are the only points of comparison currently. Though there is no discussion of the stative possession variable per se, Hickey (2007) does contain several utterances of contemporary Irish English (used to exemplify other linguistic features) that contain both the have got and got forms.
Ontario in order to ensure maximum comparability. Statistical models are maximally comparable when they are constructed with identical input parameters. The parameters (i.e., conditioning factors) that were tested across studies, and thus included in this one, are detailed in Section 7.5.3 below.

All instances of have, have got, and got, including phonetically reduced forms, with a stative possessive meaning in the present tense were extracted from the data. These verbs in negative and interrogative phrases were extracted (and are included in the token numbers in Table 7.1), but were excluded from the statistical analysis in Section 7.5.3.5. This is because variation within these contexts is confounded by differing rules in North American and British English for the use of do-support with possessive have. Furthermore, Tagliamonte and Denis (2014) do not include these contexts in their analysis of Southern Ontario communities and including them in the present analysis makes the datasets less reliably comparable. Similarly, past and future temporal reference contexts and modal constructions were not extracted because they lie outside the envelope of variation. If the forms do occur in these contexts they have a dynamic rather than stative meaning, as in (44). In addition, they were not included in previous analyses.

(44) Past *I would have got tea in my cupboard. (=stative)
Past I would have got tea at the store. (=dynamic)
Past *I would got tea in my cupboard. (=stative)
Past *I would got tea at the store. (=dynamic)
Future *I will have got tea in my cupboard. (=stative)
Future I will have got tea at the store. (=dynamic)
Future *I will got tea in my cupboard. (=stative)
Future *I will got tea at the store. (=dynamic)

7.5 Results

A total of 4,670 tokens were extracted from the four Cape Breton English corpora. There were 123 tokens of stative possessive verbs in the dialogue in the McKinnon Texts. These tokens were examined qualitatively. The pattern of variation within these 123 tokens is suggestive of the kind of variation that may have been within the community in the 1850s — namely that variation between have, have got, got was present in the community during this time. The distribution of variants in the remaining 4,547 tokens from the Storyteller, Steelworker, and Post-Industrial corpora is presented in Table 7.1.

Table 7.1: Distribution of stative possessive verbs in Cape Breton English— all present-tense contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>have</th>
<th></th>
<th>have got</th>
<th></th>
<th>got</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>Storyteller</td>
<td>89</td>
<td>47</td>
<td>75</td>
<td>40</td>
<td>24</td>
<td>13</td>
<td>188</td>
</tr>
<tr>
<td>Steelworker</td>
<td>2,564</td>
<td>76</td>
<td>590</td>
<td>18</td>
<td>202</td>
<td>6</td>
<td>3,356</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>731</td>
<td>73</td>
<td>93</td>
<td>9</td>
<td>179</td>
<td>18</td>
<td>1,003</td>
</tr>
<tr>
<td>Total</td>
<td>3,384</td>
<td>74</td>
<td>758</td>
<td>17</td>
<td>405</td>
<td>9</td>
<td>4,547</td>
</tr>
</tbody>
</table>

7.5.1 Stative possessive verbs and do-support

The nature of the have verb has been widely cited as a distinction between North American and British English (Trudgill 2010, Ch. 5; Biber et al. 1999; Quirk et al. 1985; Gramley and Pätzold 1992, among others). The distinction is between which functions of have require do-support. In North American English when have acts as a
main verb it requires the dummy auxiliary do in contexts of negation, interrogation, ellipsis and emphasis, just like any other lexical verb. When have acts as an auxiliary (e.g., to express perfective aspect), it does not require do-support, just like other auxiliary verbs.

(45) Main verb I have tea after 6 pm on workdays. (=dynamic)
    I do not have tea after 6 pm on workdays.
    *I have not tea after 6 pm on workdays.

Main verb I have tea in my cupboard. (=stative)
    I do not have tea in my cupboard. (North American English/innovative British English)
    I have not tea in my cupboard. (conservative British English)

Auxiliary verb I have drunk tea.
    *I do not have drunk tea.
    I have not drunk tea.

Auxiliary verb I have been drinking tea.
    *I don’t have been drinking tea.
    I have not been drinking tea.

In conservative British English varieties, however, main verb have only requires do-support when it has a ‘dynamic’ (receiving, taking, or experiencing) meaning, e.g., I don’t have tea after 6 p.m. on workdays, but not when it has a ‘stative’ (possessive) meaning, e.g., I haven’t any tea in my cupboard (Trudgill 2010, 110–113). Do-support for main verbs is a change that has occurred in the English language and the spreading of do-support to have as a main verb has not fully penetrated Scottish and Irish English (Trudgill 2010, 115). Furthermore, do-support began with dynamic verbs and then spread to stative verbs, so the North American system is “newer” than the British System.

Tagliamonte and Denis (2014, 99) conclude that “in a Canadian context, do-support for stative have has virtually gone to completion,” conforming to claims made by Trudgill, Nevalainen, and Wischer (2002, 6) for North America more generally. The authors report that in Toronto (based on Tagliamonte, D’Arcy, and Jankowski 2010) and southeastern Ontario, negatives and questions appear nearly categorically with have. Looking first at negatives, of the negative tokens in Tagliamonte and Denis’s (2014) data which do not use no-negation (which does not require do-support), 88 percent (N=107) use do-support. Of the remaining thirteen tokens, twelve are haven’t got, in which have acts as an auxiliary, and one is have not, showing the older, more British or more Scottish/Irish pattern.

Table 7.2: Distribution of stative possessive verbs in Cape Breton English — negative present-tense contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>no-negation</th>
<th>not-negation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>have</td>
<td>have got</td>
<td>got</td>
</tr>
<tr>
<td>Storyteller</td>
<td>3</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Steelworker</td>
<td>108</td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>19</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>130</td>
<td>21</td>
<td>36</td>
</tr>
</tbody>
</table>

1 Includes all sentential negators (e.g., no, none, no one, nothing)
2 This includes one token of do-support with bare got.

In the Cape Breton data, as in Toronto, do-support is nearly categorical. In the Storyteller, Steelworker and Post-Industrial corpora there are 563 negative stative possessive constructions, of which 187 use no-negation.
Of the 376 tokens with verbal (or constituent) negation, only 258 (61 percent) employ _do_-support, compared to 88 percent in Toronto. However, this difference is partially an artefact of the overall higher rate of _have got_ tokens in the _Storyteller_ data compared to the Toronto data. _Have got_ does not require _do_-support because _have_ inverts with the subject. Looking instead just at contexts where _do_-support is expected to occur and where _have_ inversion would be ungrammatical — that is, _have_ alone with no _got_ — only one token among Tagliamonte and Denis’s (2014) 120 verbal negation tokens (0.83 percent) was _have not_, while the rest where _do not have_. In Cape Breton a similar rate was found: 3 tokens in 261 tokens (1.1 percent). Though, of these three tokens without _do_-support, see (46) below, two unexpectedly come from younger speakers. This indicates that even though the stative possessive system has undergone change over the 20th century towards a fully-lexical _have_, outdated syntactic constructions (like _haven’t_) as well as retrogressive variants (like bare _got_) persist.

(46) a. He called his friend and then told him, “Well look, buddy, the doctors told me I’m gonna die and I _have not_ time to live like you have lived.” (Male born 1910, ST)
   b. I haven’t a flying idea. 1-, 1-, no idea what I’m going to do. (Male born 1989, PI)
   c. So [laughs] _I have not_ an idea! (Female born 1991, PI)

Table 7.3: Distribution of stative possessive verbs in Cape Breton English — interrogative present-tense contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th><em>have you?</em></th>
<th><em>do you have?</em></th>
<th><em>have you got?</em></th>
<th><em>you got?</em></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storyteller</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Steelworker</td>
<td>1</td>
<td>2</td>
<td>25</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>2</td>
<td>13</td>
<td>3</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>4</strong></td>
<td><strong>15</strong></td>
<td><strong>29</strong></td>
<td><strong>5</strong></td>
<td><strong>53</strong></td>
</tr>
</tbody>
</table>

There are 53 interrogative tokens in the _Storyteller_, _Steelworker_, and _Post-Industrial_ corpora. By far the most common construction is with _have got_ (e.g., _have you got?_), which occurs in 29 (55 percent) of the interrogative contexts. While _do_-support is present, used in 15 (28 percent) of the tokens, the numbers of constructions of lexical _have_ without _do_-support, like (47), are non-negligible (4 tokens, 8 percent), despite the low number of tokens. But, they are also from speakers born before 1942, while almost all the _do_-support tokens are from speakers born after 1960, suggesting a change in progress. There is also a non-negligible number of tokens with _got_ and no inversion (5 tokens, 9 percent). This again suggests again that both retrogressive variants (_got_) and outdated syntactic constructions (_have you?_) persist in Cape Breton English.

(47) a. You go in and say to him, “Charlie, have you anything for me today?” But remember, place is full, just whisper to him. (Male born 1895).
   b. Have you any idea what a crank shaft is? (Male born 1917)

To summarize, the use of _do_-support for main verb _have_ is a change in English that Tagliamonte and Denis (2014, 99) conclude has reached completion in Canadian English. This change has not occurred (or not occurred robustly) in conservative British varieties. The above results show that Cape Breton English has a pattern of _do_-support for _have_ that is equally as advanced along the trajectory of this change as speakers in Toronto and surrounding communities. This represents the first piece of evidence in this chapter that Cape Breton English patterns grammatically more similarly with mainstream Canadian English than Scottish English. If this is true, similar internal linguistic constraints for the use of stative possessive _have_ in affirmative declarative constructions should be found in both Cape Breton and Toronto/Southern Ontario English, because the corollary of this conclusion is that the two are experiencing parallel change. If different internal linguistic constraints for the use of stative possessive _have_ in affirmative declarative constructions are found instead, the conclusion would
be that despite the innovation in the use of *do*-support for lexical *have*, Cape Breton English is still grammatically dissimilar from mainstream Canadian English. Comparing the constraints to those for Scottish, Irish, or British English will therefore further illuminate from which source Cape Breton English originated.

### 7.5.2 Trajectories of Change

Table 7.4: Distribution of stative possessive verbs in Cape Breton English — present-tense affirmative declarative contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>have</th>
<th></th>
<th>have got</th>
<th></th>
<th>got</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>Storyteller</td>
<td>80</td>
<td>47</td>
<td>67</td>
<td>40</td>
<td>22</td>
<td>13</td>
<td>169</td>
</tr>
<tr>
<td>Steelworker</td>
<td>2,274</td>
<td>79</td>
<td>449</td>
<td>16</td>
<td>169</td>
<td>6</td>
<td>2,892</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>619</td>
<td>71</td>
<td>79</td>
<td>9</td>
<td>171</td>
<td>20</td>
<td>869</td>
</tr>
<tr>
<td>Total</td>
<td>2,973</td>
<td>76</td>
<td>595</td>
<td>15</td>
<td>362</td>
<td>9</td>
<td>3,930</td>
</tr>
</tbody>
</table>

Table 7.4 shows the overall distribution of the three stative possessive variants in affirmative declarative contexts by corpus. Overall there is 76 percent *have* use, which is about what Tagliamonte (2006c) reports for all Toronto speakers (77 percent) and Walker and Hoffman (2016) report for Toronto speakers with British/Irish ethnic backgrounds over the age of 40 (75 percent). The rates of the three variants, however, are not similar across the three Cape Breton corpora. The *Storyteller* data, which was collected earliest and also covers the oldest period in apparent time, has more *have got/got* (53 percent) than *have* (47 percent), and more *have got* (40 percent) than bare *got* (13 percent). The *Steelworker* data, which was collected in 1989-90 and contains tokens from speakers born in the 1900s–1940s, shows the highest rate of *have* use (79 percent), with *have got* used 16 percent of the time, and *got* used 6 percent of the time. In fact, the data from the *Steelworker* corpus best matches the Toronto data presented by Tagliamonte (2006c), even though the Cape Breton speakers are older, exclusively blue collar, and almost all men. The *Post-Industrial* Cape Breton data has a lower *have* rate (71 percent) than the *Steelworker* data, and the highest rate of bare *got* (20 percent).

The variable patterns across corpora are likely due to the uneven distribution of data by speakers in any decade when the data is broken down by corpus. For example Figure 7.3 shows a spike in *have got* (dashed orange line with triangles) use in the *Post-Industrial* corpus in the 1930s, but there is only only speaker born in that decade from the *Post-Industrial* corpus who used a stative possessive construction. Of her /five tokens, three are *have got*. The four tokens from one speaker born in the 1930s in the *Storyteller* corpus are, on the other hand, all *have*. In the *Steelworker* corpus there were 380 tokens from 23 speakers born in the 1930s, of which 49 (13 percent) were *have got*.

Given that the overall trajectory of change appears to be the rise of *have* and the decline of the *have got/got* forms over time across all the corpora, all three *have* corpora have been combined in Figure 7.4. The tokens from speakers born before 1900 (14 tokens) have been combined with the tokens from speakers born in the first decade of 1900s. For comparison, the relative rates of *have* use in Toronto, Ontario, and Buckie, Scotland, (Tagliamonte 2006c; 2013a) have been included in Figure 7.4. Cape Breton speakers born in the first decade of the 1900s show lower rates of *have* — and thus greater rates of *have got/got* — than observed in either Toronto.

---

5. The decline in *have* in the 1940s in the *Storyteller* corpus is due to the 4 tokens (all *have got*) from a single speaker born in 1940s.
6. The trajectories of change in Toronto and Buckie are stepwise because the data reported by Tagliamonte (2006c) and Tagliamonte (2013a) groups speakers into three age categories, e.g., >60, 60–30, <30. Figure 7.4 shows the same rate of *have* for all birth years that fall within these age categories.
Figure 7.3: Distribution of stative possessive verbs in Cape Breton English — all variants by corpus and decade of birth.

Figure 7.4: Distribution of stative possessive verbs in Cape Breton, Toronto and Buckie — all variants by decade of birth. Comparative rates for stative possessive *have* in apparent time for Toronto (2004) and Buckie (1997), as presented by Tagliamonte (2006a, Fig. 6; 2013a, Fig. 7.15) (solid black line) or Buckie (dashed black line). The oldest Cape Bretoners use *have got/got* 63 percent of the time, much higher than the 24 percent use Tagliamonte and Denis (2014) posited for Southern Ontario Loyalists.
based on Kroch (1989), and higher still than the approximately 30 percent have got/got use by Toronto speakers reported by Tagliamonte (2006c). This is a striking finding if Cape Breton English’s grammatical patterns are supposed to match Scottish English, where among the oldest Buckie speakers there is almost no have got and among all speakers there is no got at all (Tagliamonte 2013a, 149–150). Perhaps even more striking is the reverse pattern that occurs among those born in the 1930s onwards in Cape Breton who use have at nearly the same rate as Toronto speakers born in the same years. Rather than matching the pattern found in the Scottish data (or any British or Irish data on stative possessives), Cape Breton English is progressing through apparent time in parallel with Toronto English.

7.5.3 Linguistic Conditioning

It is conceivable that Cape Breton originally inherited a stative possessive system with no have got/got (e.g., a Scottish/Irish system) and then acquired have got/got via diffusion from Inland Canada sometime in the 19th century. This would align with the notion that Cape Breton is a relic speech enclave quickly converging with inland Canadian norms. While this is highly unlikely given the actual history of English in Cape Breton, statistical confirmation that the internal linguistic constraints for Cape Breton stative possessives match those of Ontario will provide validation for rejecting Ontario-to-Cape Breton diffusion of the retrogressive have got/got forms.

7.5.3.1 Subject Type

One hypothesis for the development of have got is that when have and has began to weaken to [v] and [z] in the 16th century, the door opened for got to be inserted as a way for speakers to more overtly express the subject-verb relationship (Crowell 1955, 283). Later, in the 1800s weak have/has is completely deleted, leaving bare got. Evidence for the validity of this hypothesis would be an early favouring of have got/got with pronominal subjects, the environment in which the contraction of have occurred first. In U.K. varieties the have got/got are favoured with pronominal subjects and disfavoured by NP subjects (Tagliamonte 2003; 2013a). In southern British communities where have got is the dominant form, the effect of subject type, though present, has weakened considerably (Tagliamonte 2013a, 157). In Canada, on the other hand, a more complex system has been found, whereby NPs, third-person plural and first person pronouns favour have, while second person and third person singular pronouns favour have got/got (Tagliamonte, D’Arcy, and Jankowski 2010; Tagliamonte and Denis 2014). These different subject types are exemplified in (48), and the distribution of stative possessive tokens in the Cape Breton data for each of these subject types are displayed in Figure 7.5. In Buckie, Scotland where have got/got is incipient, Tagliamonte (2013a) does not find a significant distinction between subject types. If Crowell’s (1955) view that pronouns were the entry point of have got/got grammaticalization is correct, the lack of a subject type constraint in Buckie suggests the have got/got forms were not a parallel independent development in Buckie, but instead a diffused change. Conversely, in the two Irish communities, Cullybackey and Portavogie, where have got/got is also incipient, have got/got forms only occur with pronominal subjects.

(48) NP Have you got little jackets for them? Gary has a waterproof one for Spike. (Female born 1927, PI)
3rd Person Plural They have a labour pool there and they have all these different sections. (Male born 1912, SW)
1st Person We have a volunteer program. We have a social worker. We have a chaplain. We have six physicians on our team. And the physician and the nurse work very closely together. One day a week, I have a physician who comes out into the community with me. If I have patients whom I feel are starting to fail, I will take a physician to the home and we’ll put all kinds of other things in place, in terms of treating their symptoms or trying to alleviate some of their symptoms. (Female born 1946, PI)
2nd Person When you got an old house you’re forever putting money in it. (Female born 1940, PI)
3rd Person Singular  I envisioned him as a larger man. He’s got a great big head, but a little body. (Male born 1991, PI)

![Distribution of stative possessive verbs by subject type in Cape Breton English](image)

Figure 7.5: Distribution of stative possessive verbs in Cape Breton English — all variants by subject type.

Testing this subject type constraint not only affords the opportunity to provide corroborating (or challenging) evidence for Crowell’s (1955) hypothesis about the development of have got/got, but it also will provide evidence that can be compared to the divergent Ontario and British stative possession systems.

### 7.5.3.2 Object Type

Jespersen (1961b) observed that when have got entered the stative possessive system, it was first used with physical concrete objects. Studies of the stative possessive system have found that object type has been a consistent trend across communities, and within communities as either have or have got increases in use over time. Kroch (1989), Tagliamonte (2013a) and Tagliamonte and Denis (2014) report that concrete objects, as opposed to abstract objects, significantly favour the use of have got/got in British, Scottish, Irish and Ontario English, though Quinn (2004) finds no significant difference between the two object types for New Zealand English.

(49) **Abstract** I guess I can say the [steel] plant has a social conscience. (Male born 1940, SW)
Concrete Well, I mean, no, you’ve got those ornaments, and Poppy’s got his cars. (Female born 1989, PI)

### 7.5.3.3 Subject Reference

In Toronto, Tagliamonte, D’Arcy, and Jankowski (2010) find a significant difference in have use with generic subjects, which favour have, and specific subjects, which favour have got/got. This constraint was not shared by speakers analyzed by Tagliamonte and Denis (2014) in the three outlying southern Ontario communities, nor was it found (or tested) in U.K. varieties.

(50)** Generic** You know, [now] you have Point Edward, and you have the Industrial Park, and you have the Coast Guard College, you have a developed college out the highway out there. (Male born 1925, PI)
Specific And I had another young fellow in particular, he [was] with me for about a year, and I said, “Why don’t you fly your fanny out of here?” I said, “B’y, now, you’ve got an education,” I said, “You’ve got some brains.” And now he is an RCMP officer down in Newfoundland. (Male born 1942, PI)
In their analysis of Toronto speakers from varying ethnic backgrounds, Walker and Hoffman (2016) find the subject of the possessive verb to be a significant predictor of have vs. have got/got; however, the authors categorize subjects differently than other studies. Walker and Hoffman divide subjects into generic pronouns, other pronouns and NPs — combining both subject reference and subject type together. They find that NP subjects favour have, generic pronouns disfavour have, and other pronouns neither favour nor disfavour have. This finding is somewhat at odds with Tagliamonte, D’Arcy, and Jankowski (2010), who find that, when considered as part of separate factor groups, both generic subjects and NPs favour have. As a greater number of studies of stative possessives follow the coding strategy used by Tagliamonte, D’Arcy, and Jankowski (2010), this is the coding strategy that will be showcased in the analysis below.  

7.5.3.4 Sex

Tagliamonte, D’Arcy, and Jankowski (2010) and Tagliamonte and Denis (2014) find that across Ontario young women lead in the use of have — conforming with Labov’s (2001, 266, and elsewhere) observation that women generally lead in the use of incoming forms. While Tagliamonte (2013a) shows that older speakers use more have, as have got/got is the innovation in Irish and Scottish English – which suggests women lead in the forms’ use – the author does not comment on which sex uses each variant most frequently. (Though Tagliamonte 2012a reports that stigmatized bare got is used more frequently by men in the British Isles generally.)

While Walker and Hoffman (2016) find a significant difference between older and younger speakers with British/Irish ethnic origin, the authors do not find a significant difference between sexes among the same speakers.

Table 7.5: Linguistic and social constraints governing the use of have for stative possession across varieties.

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Toronto¹</th>
<th>Southern Ontario²</th>
<th>Scotland³</th>
<th>Ireland³</th>
<th>Northern England⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Type</td>
<td>NP&gt;3Pl&gt;1&gt;2&gt;3Sg</td>
<td>NP&gt;3Pl&gt;1&gt;2&gt;3Sg</td>
<td>No Difference</td>
<td>NP&gt;Pro</td>
<td>NP&gt;Pro</td>
</tr>
<tr>
<td>Object Type</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
</tr>
<tr>
<td>Subject Reference</td>
<td>Generic&gt;Specific</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
<td>Concrete&gt;Abstract</td>
</tr>
<tr>
<td>Age</td>
<td>Young&gt;Old</td>
<td>Young&gt;Old</td>
<td>Old&gt;Young</td>
<td>Old&gt;Young</td>
<td>Old&gt;Young</td>
</tr>
<tr>
<td>Sex</td>
<td>Women&gt;Men</td>
<td>Women&gt;Men</td>
<td>Women&gt;Men</td>
<td>Women&gt;Men</td>
<td>Women&gt;Men</td>
</tr>
</tbody>
</table>

¹ Tagliamonte, D’Arcy, and Jankowski (2010) ² Tagliamonte and Denis (2014) ³ Tagliamonte (2013a) ⁴ Tagliamonte (2003; 2013a)

7.5.3.5 Logistic Regression Analysis

To determine which, if any, of the factors in Table 7.5 constrain the variation between have and have got/got in Cape Breton English, mixed effects logistic regression modelling, as implemented by the lme4 package in R (Bates, Maechler, et al. 2015; R Core Team 2015)³ was used, with the binary selection of have versus have got/got as the dependent variable. The model’s parameters (e.g., factors/effects) were the main fixed effects of subject type, subject reference, object type, year of birth (centred) and sex, plus the random intercept of speaker, which eliminates possible distortion caused by the uneven contribution of tokens per speaker to the data, and individual speakers’ variable propensity for using have. This full model was subjected to a Wald χ² test (Fox and
Weisberg 2011), which identified the factors that most contributed to a statistical, explanatory understanding of the data (see Footnote 6 on p. 62).

Table 7.6: Analysis of deviance — Wald $\chi^2$ test for full model. *Have* as a stative possession verb in Cape Breton English.

<table>
<thead>
<tr>
<th>Parameter (factor)</th>
<th>df</th>
<th>$\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Type</td>
<td>4</td>
<td>78.16</td>
<td>$4.3 \times 10^{-10}$ ***</td>
</tr>
<tr>
<td>Corpus</td>
<td>2</td>
<td>42.37</td>
<td>$6.3 \times 10^{-08}$ ***</td>
</tr>
<tr>
<td>Year of Birth, centred</td>
<td>1</td>
<td>34.11</td>
<td>$5.2 \times 10^{-09}$ ***</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>11.14</td>
<td>$8.5 \times 10^{-04}$ ***</td>
</tr>
<tr>
<td>Object Type</td>
<td>1</td>
<td>5.02</td>
<td>$4.00 \times 10^{-02}$ *</td>
</tr>
<tr>
<td>Subject Reference</td>
<td>1</td>
<td>1.85</td>
<td>$1.70 \times 10^{-04}$</td>
</tr>
</tbody>
</table>

* * * $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

The $\chi^2$ value is the test statistic. Degrees of freedom (df) is the number levels for a given parameter minus 1. The p-value is determined by comparing the test statistic and the df to the $\chi^2$ distribution.

Table 7.6 lists the results from the Wald $\chi^2$ test, which show that that subject type — coded here to test the Ontario pattern in Table 7.5 — corpus, year of birth, sex, and object type are significant predictors of the variation, while subject reference is not. Subject type, which has a much larger $\chi^2$ value, has a greater magnitude of effect relative to other factors. Corpus and year of birth are both significant predictors; however, as outlined in Chapter 5, these two parameters are not independent of each other. Furthermore, given the dramatic change over time observed in Figure 7.4, before proceeding with further statistical analyses, the social constraints of corpus, sex and age were analyzed using a conditional inference recursive partitioning tree (see Section 5.1 p. 44) and the data was divided into three groups. Figure 7.6 shows that the patterns observed in the distributional data above are in fact significant.

9 The topmost node in the tree is the shock point of the birth year 1930. Speakers born in 1930 or before use significantly less *have* and more *have got/got* than speakers born after 1930 ($p<0.001$). The next node in the tree is the shock point of birth year 1907; speakers born in or before 1907 use more *have got/got* than those born later ($p<0.001$), and for this group there is also a significant difference between men and women ($p=0.002$), with men using more *have got/got* and women using more *have*. After determining these shock points and partitioning the data, separate Wald $\chi^2$ tests were performed (Table 7.7) on each partition.

The Wald $\chi^2$ test on the partitioned data shows that, even though year of birth is selected as a highly explanatory parameter in the full model for all speakers, within each group of speakers ($\leq 1907$, 1908–1930, $\geq 1931$) year of birth is not significant.

For the $\leq 1907$ speakers, sex is the only significant factor. For subject reference, none of the six generic subject tokens occurred with *have* while 39 percent of the specific subject tokens did. Given that generic subjects were categorically *have got/got* this parameter was not included in the regression model for the $\leq 1907$ speakers. Subject reference might seem like an important linguistic constraint for the $\leq 1907$ speakers; however, all six of the generic subject tokens were from male speakers. In the remaining 77 specific subject tokens, 44 percent of the male tokens occurred with *have*, while 73 percent of the female tokens occurred with *have*. In Tagliamonte, D’Arcy, and Jankowski’s (2010) analysis of Toronto speech, generic subjects conversely favoured *have*, and

---

9. Corpus was not included in the analysis shown for simplicity. See Figure A.1 on p. 199 for a conditional inference recursive tree analysis that includes corpus and that also identifies 1907 and 1930 as shock points.
Toronto is the only community in which this constraint was found to be operational. For these reasons, it was concluded that the subject reference effect patterning for \( \leq 1907 \) speakers was instead epiphenomenal to the sex effect.

For the speakers born between 1908 and 1930, subject type is the only significant factor. The significance of subject type is shared by those born after 1930, but for this group object type is also a significant factor. Based on these tests, the Cape Breton data shows patterns of variation that are more like Ontario than Scotland. In Buckie there is not a significant difference for subject type (Tagliamonte 2013a, 155), yet in the bulk of the Cape Breton data this parameter provides the most explanatory understanding of the variation — as it does in the Ontario communities. Like Buckie, the oldest Cape Breton speakers do not show a significant effect for subject type, but unlike Buckie speakers, this group uses *have got/got* frequently — and categorically with generic subjects.

Cape Breton speakers born after 1907 differ from Toronto speakers insofar as subject reference is not selected as a significant parameter; however, Toronto is the only community in which this factor seems to be operational. The three outlying communities in Southern Ontario studied by Tagliamonte and Denis (2014) did not show evidence of this constraint, just like Cape Breton.

Table 7.8 presents the results of a mixed effects logistic regression analysis of the most parsimonious model for each Cape Breton age groups based on the Wald \( \chi^2 \) tests presented in Table 7.7. The results of these analyses
Table 7.7: Analysis of deviance — Wald χ² test for full model with data partitioned by year of birth. *Have* as a stative possession verb in Cape Breton English.

<table>
<thead>
<tr>
<th>Parameter (Factor Group)</th>
<th>df</th>
<th>≤1907 χ²</th>
<th>p-value</th>
<th>1908–1930 χ²</th>
<th>p-value</th>
<th>≥1931 χ²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Type</td>
<td>4</td>
<td>3.86</td>
<td>4.26 × 10⁻¹¹</td>
<td>42.08</td>
<td>1.60 × 10⁻⁸</td>
<td>***</td>
<td>48.54</td>
</tr>
<tr>
<td>Object Type</td>
<td>1</td>
<td>0.04</td>
<td>8.39 × 10⁻¹¹</td>
<td>0.07</td>
<td>8.00 × 10⁻¹</td>
<td>***</td>
<td>0.07</td>
</tr>
<tr>
<td>Subject Reference</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1.33</td>
<td>2.50 × 10⁻¹</td>
<td>***</td>
<td>1.73</td>
</tr>
<tr>
<td>Year of Birth, centred</td>
<td>1</td>
<td>0.07</td>
<td>7.94 × 10⁻¹¹</td>
<td>0.07</td>
<td>7.90 × 10⁻¹</td>
<td>***</td>
<td>1.61</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>4.66</td>
<td>3.10 × 10⁻²</td>
<td>*</td>
<td>0.11</td>
<td>7.40 × 10⁻¹</td>
<td>*</td>
</tr>
</tbody>
</table>

* * * p < 0.001, ** p < 0.01, * p < 0.05

The χ² value is the test statistic. Degrees of freedom (df) is the number levels for a given parameter minus 1. The p-value is determined by comparing the test statistic and the df to the χ² distribution.

1 Subject Reference was omitted from the ≤1907 group because none of the six generic subject tokens occurred with *have*.

are presented in a table formatted akin to those for reporting results from the multivariate analysis tool *Goldvarb* (Sankoff, Tagliamonte, and Smith 2015) as these *Goldvarb*-style tables are more common in variationist sociolinguistic literature.10 The factor weights for the non-significant fixed effects come from the full models, while the factor weights for the significant fixed effects, the inputs, the Akaike information criterion (aic) values, and the random effect variances come from the most parsimonious models (i.e., models built with only significant factor groups, following Bates, Kleigl, et al. 2015). A secondary analysis of the most parsimonious models testing the contrast between levels of each parameter (e.g., factors in each factor group), rather than between each level and the mean, was completed as well. For factor groups with more than two factors (here subject type), a horizontal line is used to indicate where significant contrasts exist.

Table 7.8 shows that for speakers born in 1908-1930 and from 1930 onwards subject type is significant and there is the same ordering of subject type factors as reported for Ontario by Tagliamonte, D’Arcy, and Jankowski (2010) and Tagliamonte and Denis (2014). In both the 1908-1930 and ≥1931 groups, NPs favour *have* use the most. For these speakers the difference between NP subjects and third person plural pronominal subjects is not significant, indicating that *have* is more or less equally likely with either subject type.11 As subject types progress along the cline of NP>3rd Person Plural>1st Person>2nd Person>3rd Person Singular, the likelihood of *have* increasingly lowers. Cape Breton shares the same very specific Canadian subject pattern with Ontario, and aligns neither with Scotland (i.e., Buckie), where subject type is not a significant predictor, nor Ireland/Northern England, where the relevant subject distinction is between NPs and all pronominal subjects (which show no internal differences). Cape Breton speakers born after 1930 also match Ontario English speakers (though also Scottish and Northern British English speakers) of the same age cohort in showing marginal preference for *have* with abstract objects. These findings are statistical validation that the Cape Breton stative possessive system matches that of Ontario, and provides strong evidence that the stative possession system in Cape Breton and Ontario were both inherited through transmission from some earlier common variety.

11. Alternative models were created in which singular and plural NP subjects were treated separately; however, neither type of NP was significantly different from each other or plural third person pronominal subjects.
Table 7.8: Mixed effects logistic regression — contribution of factors to the probability of *have* for expressing stative possession across three age groups in Cape Breton English.

<table>
<thead>
<tr>
<th>Stative possessive <em>have</em> in Cape Breton English</th>
<th>≤ 1907</th>
<th>1908–1930</th>
<th>≥ 1931</th>
</tr>
</thead>
<tbody>
<tr>
<td>aic</td>
<td>99</td>
<td>1411</td>
<td>2231</td>
</tr>
<tr>
<td>input</td>
<td>.49</td>
<td>.71</td>
<td>.88</td>
</tr>
<tr>
<td>total N</td>
<td>83</td>
<td>1221</td>
<td>2626</td>
</tr>
</tbody>
</table>

### Fixed Effects

**Subject Type**

<table>
<thead>
<tr>
<th></th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun Phrase</td>
<td>.30</td>
<td>8</td>
<td>.66</td>
<td>80</td>
<td>.64</td>
<td>87</td>
</tr>
<tr>
<td>3rd Person (plural)</td>
<td>.30</td>
<td>29</td>
<td>.62</td>
<td>78</td>
<td>.65</td>
<td>89</td>
</tr>
<tr>
<td>1st Person</td>
<td>.71</td>
<td>46</td>
<td>.48</td>
<td>69</td>
<td>.51</td>
<td>81</td>
</tr>
<tr>
<td>2nd Person</td>
<td>.53</td>
<td>31</td>
<td>.47</td>
<td>65</td>
<td>.36</td>
<td>74</td>
</tr>
<tr>
<td>3rd Person (singular)</td>
<td>.44</td>
<td>44</td>
<td>.13</td>
<td>44</td>
<td>.34</td>
<td>67</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Object Type**

<table>
<thead>
<tr>
<th></th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>.51</td>
<td>42</td>
<td>.43</td>
<td>65</td>
<td>.55</td>
<td>82</td>
</tr>
<tr>
<td>Concrete</td>
<td>.49</td>
<td>34</td>
<td>.57</td>
<td>69</td>
<td>.45</td>
<td>80</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Subject Reference**

<table>
<thead>
<tr>
<th></th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>[–]</td>
<td>0</td>
<td>[–]</td>
<td>6</td>
<td>[–]</td>
<td>66</td>
</tr>
<tr>
<td>Specific</td>
<td>[–]</td>
<td>39</td>
<td>[–]</td>
<td>77</td>
<td>[–]</td>
<td>69</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sex**

<table>
<thead>
<tr>
<th></th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>.75</td>
<td>73</td>
<td>.52</td>
<td>77</td>
<td>.48</td>
<td>77</td>
</tr>
<tr>
<td>Male</td>
<td>.25</td>
<td>28</td>
<td>.48</td>
<td>67</td>
<td>.52</td>
<td>68</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Year of Birth**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>continuous factor</td>
<td>[–]</td>
<td>[–]</td>
<td>[–]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Random Effects

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
<th>N</th>
<th>Variance</th>
<th>N</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEAKER</td>
<td>1.41</td>
<td>29</td>
<td>1.04</td>
<td>95</td>
<td>1.90</td>
<td>154</td>
</tr>
</tbody>
</table>

Sum contrast coding.

Aic, input, and factor weights for significant factor groups derived from most parsimonious model. Factor weights for non-significant factor groups derived from full model.

Factor weights converted from log-odds. $p = e^{\frac{\sigma^2}{1+\sigma^2}}$

Factor weights with non-significant contrasts indicated with [ ].

Correlation of fixed effects: ≥1931 $r < |0.068|$

Additional treatment contrast coding used to test orthogonal contrasts for parameters (factor groups) with more than two levels (factors) using most parsimonious model.

Significant contrasts indicated with horizontal lines.
7.6 Discussion

7.6.1 have/have got/got

With respect to the change of stative possessives in Cape Breton, the preliminary evidence suggests transmission — that is, the patterns of have and have got/got usage in Cape Breton reflect a trajectory of change that has a common origin with changes that are occurring elsewhere in Canada. The use of have got in the McKinnon Texts indicates have got was part of the inherited linguistic system in Cape Breton English. In the first decades of the 20th century have got/got was a robust form, as evidenced by the have got/got use of Cape Breton speakers born before 1930 — more robust, in fact, than for the oldest speakers examined by Tagliamonte and Denis (2014) born in the same era. Among those born after 1930 there are striking similarities between Cape Breton Inland Canada, not just for which variant is most common, but also the ways in which the variants are used (i.e., internal constraints). Diffusion from Inland Canadian English is unlikely in this case considering there is no apparent loss of linguistic conditioning between Cape Bretoners and Ontario speakers and because, for the most part, after 1908 Cape Breton leads Ontario in the use of have.

It is important to note that the heaviest have got/got users are not just the oldest speakers in the Cape Breton data, but also the speakers who come from rural (= formerly Gaelic-speaking) communities. The language shift that took place in these communities resulted in rural speakers using a high frequency of a linguistic feature with a direct provenance in Loyalist speech. Given that the language shift was engineered by the provincial government via English-exclusive education, the agents of change involved in shifting the community language were children. This is perhaps why the internal linguistic constraints governing the variation between forms so closely matches those of other Loyalist communities, including the nearby Loyalist-settled towns. The large number of 19th-century Yankee tourists to rural Cape Breton undoubtedly also reinforced the overt prestige of specifically Yankee English — as did the large number of young Cape Breton Gaels (especially women) traveling back and forth to the more metropolitan “Boston States” (New England) for (often domestic) employment (e.g., Aberdeen 1960, 229). The close alignment between the descendant dialects of coastal Loyalists (Cape Bretoners) and inland Loyalists (Ontarians) suggests that the variation between stative possessive forms pre-dates the divergence of these two dialects.

Looking at the more recent data, of note are the similar findings reported by Woods (1999, 190-191) who, in 1979, found “a strong generational gap” between survey respondents born before and after 1937 in Ottawa. When asked the question If you needed a match, what would you ask your friend: ______ a match?, respondents born before 1937 most often selected Have you got…? while those born after 1937 most often selected Do you have…? Younger respondents with newer Canadian backgrounds (this is how Woods refers to new Canadians or first-generation Canadians) were even more likely to select the latter. In his study Woods, maybe erroneously, labels Have you got…? as Canadian and Do you have…? as American, though his labelling would conform to the idea that have got/got was much more common in Canada generally before WWII (cf. Fowler 1983, 24). Woods (1999, 52) hypothesized many variables would show divergent patterns for those born before 1937 and those born

12. Neither Tagliamonte, D’Arcy, and Jankowski (2010) nor Tagliamonte and Denis (2014) included speakers born before 1907. Speakers born before 1907 in Ontario, for whom the have got/got forms were likely more prevalent, may show the same linguistic conditioning as Cape Breton. Furthermore, in British communities where have got is increasing the effect of subject type is weakening or has levelled. For speakers born before 1907 have is a minority form, so it is perhaps unsurprising that it does not have a significant subject type effect.

13. In his 1970 book, Speaking Canadian English, Orkin calls interrogative Have you…? and Have you got…? the unquestionable “shibboleth of Canadian syntax,” (Orkin 1970, 155). Citing a questionnaire study (it is unclear which study Orkin is referring to, though it is likely Avis (1955) and Hamilton (1955)), Orkin (1970, 156) says that the American Do you have…? “enjoyed almost no currency in Ontario, only 10 percent of the limited number of persons tested using the American form, while 3 percent used both forms interchangeably.” In Montreal, on the other hand, “do you have was offered by 35 percent of informants, the remaining speakers divided their allegiance between Have you (30 per cent), and have you got (35 percent), although the latter was frowned upon by English teachers because of the redundant “got”, (Orkin 1970, 156).
afterwards (The Ottawa Survey of Canadian English was actually designed to test this idea) because of the rapid socio-demographic and economic changes that occurred in Ottawa (like many communities) during and after WWII. According to Woods’s (1999, 276) results this gap was most striking for grammatical variables. Authors studying other English-speaking communities (e.g., Bailey et al. 1996; Beal and Grant 2006) have categorized WWII as a catastrophic event that caused “significant linguistic consequences” (Bailey et al. 1996, 449), especially among those born during and after the war.

Though the relevant division in apparent time is a few years earlier, generally the generation that came of age before WWII in Cape Breton (≤ 1907) differ from those who came of age during it, and those that came of age afterwards. Among those that came of age during the interwar and WWII era (1908–1930) we see a near alignment with mainstream Canadian English norms. Among those who were born later (≥ 1930) Cape Breton and Ontario are indistinguishable.

As stated previously, Cape Breton English has been described as a relic dialect rapidly converging with mainstream Canadian norms. The stative possessive system in Cape Breton does show a strong alignment with that of Inland Canada, which would be indicative of this rapid convergence — but were the two varieties divergent to begin with? And is this divergence in the direction we would expect? If the difference in founding conditions between Cape Breton and Southern Ontario is that there was a much greater proportion of Scottish/Irish immigrants to Cape Breton, conservative Cape Breton English should be more Scottish/Irish-like. However, this is not the case. Conservative Cape Breton English — the speech of older Cape Bretoners, and especially rural older Cape Bretoners in the Storyteller corpus — is the most non-Scottish/Irish because conservative Scottish/Irish English has almost no have got/got use, while conservative Cape Bretoners have robust have/have got/got variation. While the Scottish/Irish influence on Cape Breton English is non-negligible — the presence of the after-perfect, ingressive “yes” and “no”, and slit-fricative /t/ and /d/, for example, can only derive from Celtic origins — the importance of the initial Loyalist input cannot be downplayed.

The remarkable similarity between the Cape Breton and Ontario stative position re-positions Cape Breton English not as a relic speech island, or a transported Scottish English, but rather the most easterly form of Loyalist-origin Canadian English. Studying trajectories of language change in Cape Breton English can therefore provide insight into the history of Canadian English and its homogenization, especially in communities like Cape Breton with mixed-origin founders.

### 7.6.2 Have got/got

While the rise in have use over the 20th century does conform with Keifte and Kay-Raining Bird’s (2010, 62) claim that Cape Breton speakers are quickly converging with the prestigious Inland Canadian dialect, among speakers born after 1960 there seems to be an additional change occurring in contexts where have is not used. Figure 7.4 shows that prior to the 1960s in apparent time, the most common of the two have got/got forms is have got, but after the 1960s the most common of the two is got. Of note, 1960 is the first year in the dataset in which all speakers come from the Post-Industrial corpus only.

A conditional inference tree, Figure 7.7, testing the three linguistic predictors, year of birth, sex, and corpus identified that the Post-Industrial corpus, overall, contains significantly more got relative to have got than the Storyteller and Steelworker corpora ($p < 0.001$). This is perhaps unsurprising given the results reported in Tables 7.4 and 7.1. It also identified that within the Post-Industrial corpus there is a significant ($p < 0.001$) contrast between third person pronominal subjects and first/second person pronominal and other subjects. In the Storyteller and Steelworker corpora, however, the contrast is instead between third-person singular subjects, and all other subject types.
Figure 7.7: Conditional inference recursive partitioning tree — got vs. have got in Cape Breton English. Parameters included: subject type, object type, subject reference, corpus, year of birth and sex.
Table 7.9: Stative possessive got vs. have got in Cape Breton and Toronto English, affirmative polarity present tense only.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Post-Industrial</th>
<th>Storyteller/Steelworker</th>
<th>Toronto¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>% got</td>
<td>N</td>
</tr>
<tr>
<td>2nd Person</td>
<td>46</td>
<td>83</td>
<td>249</td>
</tr>
<tr>
<td>1st Person</td>
<td>110</td>
<td>81</td>
<td>201</td>
</tr>
<tr>
<td>3rd Person (plural)</td>
<td>19</td>
<td>47</td>
<td>91</td>
</tr>
<tr>
<td>Noun Phrase</td>
<td>22</td>
<td>63</td>
<td>56</td>
</tr>
<tr>
<td>3rd Person (singular)</td>
<td>53</td>
<td>39</td>
<td>110</td>
</tr>
<tr>
<td>Total</td>
<td>250</td>
<td>68</td>
<td>707</td>
</tr>
</tbody>
</table>

¹ Tagliamonte, D’Arcy, and Jankowski (2010, 165)

The pattern of variation for got vs. have got in both the Post-Industrial and Storyteller/Steelworker groups matches that of Southern Ontario English and other dialects of English whereby third-person singular subjects resist bare got (Tagliamonte, D’Arcy, and Jankowski, 2010, 160, see Table 7.9) relative to other subject types. That said, bare got with singular third person pronominal subjects was not completely absent in any of the datasets.

(51) Storyteller I said, “Boy, he got two arms and two legs and he can shovel - that’s all I want. (Male born 1914)
Steelworker But this is the kind of stuff that takes place in the steel plant. And there is one type of guy he got his own opinions and the devil in hell is not going to change him. He’s either anti-company or anti-boss, or he’s a boss and he is anti-man and you’re not going to change him.
Post-Industrial She got too many grandchildren! (Female born 1965)

When NP subjects were divided based on number, singular NPs from the Storyteller/Steelworker data behaved more like third person singular subjects (16 percent got), while plural NP subjects were more like other subject types (26 percent got). This suggests that the relevant distinction for got versus have got within this dataset aligns with the pattern of (partial) suppletion for the verb have. In contexts where either have got or got are selected over have, got is most common when have got would otherwise be realized as have or ’ve (e.g., non-third person singular). In contexts where have got is required to be realized as has or ’s (e.g., third person singular), has got or ’s got is more common. Tagliamonte, D’Arcy, and Jankowski (2010) indicate this is also likely the case in Toronto. The authors report that, as got arises from the contraction of have, it is likely that the contraction of have to ’ve and then zero is a change that is related and parallel to but not necessarily at the same stage as the contraction of has to ’s and then zero.

While the overall rate and linguistic conditioning of got versus have got in the Storyteller/Steelworker data is comparable to the rate in Toronto, the rate in the Post-Industrial corpus is about three times as high. Furthermore, over both real and apparent time the linguistic conditioning for when got is selected over have got changes. What was likely a morphological constraint based on (partial) suppletion instead becomes one based on subject type. Third person pronominal subjects, whether singular or plural, resist bare got (p < 0.001) and in contexts where the subject favours got, the variant is significantly more favoured by males (p = 0.005).

There are both social and linguistic factors driving the increase in likelihood of got over have got as the main secondary variant. Linguistically, the have got variant challenges the status of possessive have as a full lexical verb. A system that only includes possessive have and possessive bare got is one in which the status of
possession have as a lexical verb is unambiguous. It also permits have, in its function as an auxiliary for marking perfect aspect in English, to be unambiguously functional. A change towards a system in which have and got are the only options is consistent with the change of stative have to a full lexical verb with do-support. In fact, the pattern observed in the Cape Breton data is perhaps the best evidence presented so far that shows this interrelated pattern (see also D’Arcy 2015).

Consider the following passage from an interview with a male speaker born in 1989:

(52) Speaker: I don’t know how much you know about coming over Carpenter’s Hill in River Ryan, when you’re going towards New Waterford.
Interviewer: Oh, yeah, yeah, yeah, yeah.
Speaker: You come over and then you got that turn, right?
Interviewer: Yeah.
Speaker: But then you got the turn for the other way. And I was following a buddy of mine, and came up there. And I was flying. Like I was doing maybe 70 or 80 going around, and that’s pretty...
Interviewer: Fast for there.
Speaker: …fast for that turn, right? So I slowed down a little but, but I’m. I probably did the turn at about, like, 50 or 60, in a ‘92 Nissan Sentra, tires squealing. [makes police siren noises] He’s like, “I’m pulling you over for reckless driving.” I said, “Jesus officer—.” I wasn’t wearing shoes. Like, I’m hiding my feet, like underneath, and, “Oh Jesus!”, and I don’t even know if I had a shirt on. I had a wife beater on I think, maybe, right?
Interviewer: Yeah.
Speaker: And ah, he’s like “You know a girl almost died the other day in a car accident?” And it was [NAME] that he was talking about. He was like, he’s like “I’m the lead investigator on that.” I’m like, “Ah Jesus!”, right, and that. I’m like “Ah, shit. Now I feel bad,” right?
Speaker: Yeah.
Speaker: So, he goes back. He says, “You got your license?” I looked in my wallet, I didn’t have my license on me. He’s like, “So, what’s your name?” He’s like, “Do you have a license?” I was like, “Yes.” I was like, “Yes, I HAVE a license.”

Twice the speaker uses got with generic subject you to introduce existential clauses; this the most common use of generic subjects with have got/got forms in the data and perhaps the high frequency use of you got… or you’ve got… for existential clauses rather than there is/are… is particular to Cape Breton — explaining the lack of a subject reference constraint in the data compared to Toronto. These two tokens of got do not speak to the form’s indexical use. The other token does. When the speaker employs the got form for the police officer’s first question about the license, the speaker lowers the pitch of his voice and slows his speech rate — both stereotypical features of masculinity (see, for example McConnell-Ginet 1983). The second time the police officer asks about the license the speaker does not lower his pitch or slow his speech rate, but puts on a more questioning and exasperated tone. In this quotation he employs the have form. The speaker also employs the have form in his polite, albeit coy, response to the officer. Hypothetically, this variant use of stative possessive forms could indicate that the speaker is using the forms to highlight an aspectual contrast: got your license meaning ‘having the license on me at that moment’; and have a license meaning ‘having a license generally.’ On the other hand, the fact that the got form is used as part of the performance of a masculine or “tough guy” persona suggests instead that form is indexically linked to conceptions of “toughness”, and “masculinity”. These qualities overlap with the working-class, traditional Cape Breton persona (Robertson 1991), so the use of got in this story by the speaker is a creative and stylistic way to situate the police officer in a specific local space and ascribe to the officer a set of characteristics and ideologies that the interviewer (also a local) would be familiar with. Older Cape Breton speakers show robust variation between have, have got, and got, but the greater stigmatization of got off-island likely means the bare got form is the one with stronger indexical connections to local emblematic identities.

The use of bare got shows a similar pattern to the use of the traditional fricated realization of word-final (t.d) (e.g.[!]0]) among the exact same young speakers. This feature, which is unquestionably an Irish/Gaelic or

14. The data from Toronto does not comment on the existential use of generic subjects and have got/got so any difference here is speculative.
Irish/Scottish English substrate feature in Cape Breton English,\(^{15}\) shows stability in apparent time across the 20th century in its use relative to other (t,d) variants (e.g., deletion, stop plus aspiration, glottal stop) through apparent time in the Post-Industrial corpus. As a stable sociolinguistic variable, a given token is more likely to be fricated if the speaker is a male (following Labov 2001). Among the youngest male speakers, however, there is an increase in the likelihood of this traditional, male feature. Which, again, may be indicative of a general trend among the young men in the corpus to reject external linguistic norms and embrace linguistic features of their forefathers.

In her 2012 plenary address at the New Ways of Analyzing Variation conference, Tagliamonte (2012a) offered

---

\(^{15}\) The frication of word-final /t/ — and /d/ or other plosives — is a heritage effect that is well-attested in many Irish or Irish English-influenced varieties of English, including the dialects of southern Ireland \((\text{inter alia} \text{ Hickey 1984; 1986; 1996; 1999; 2004; Jones and Llamas 2008; Ó Baoill 1990; 1997; Pandeli et al. 1997; Wells 1982a); Liverpool, Newcastle, Tyneside, and Middlesbrough, England. (Docherty and Foulkes 1996; Jones and Llamas 2008; Knowles 1973; 1978; Marotta and Barth 2005; Sangster 2001; Watson 2002; 2007a; 2007b; Watt and Allen 2003; Wells 1982a); the Shetland Islands, off northern Scotland (Melchers and Sundkvist 2010); Sydney, Adelaide, Canberra, and Melbourne, Australia (Haslerud 1995; Jones and McDougall 2009; Tolfree 1996; 2001); and southeastern Newfoundland, as well as Halifax and Cape Breton Island, Nova Scotia, Canada (Clarke 1986; 2010; Gardner 2013a; Keifte and Kay-Raining Bird 2010; Parris 2009; Power 2010; Seary, Story, and Kirwin 1968).\)
a synthesis of the dual motivations of change within the stative possessive system. She argued that bare got in the U.K. both embodied the linguistic evolution of the stative possession system — by being favoured in specific contexts, e.g., with first/second person pronominal subjects and with concrete objects — while at the same time indexing emblematic identities. These identities include authentic, specifically local identities. The same is likely occurring among young speakers in Cape Breton.

7.7 Conclusion

This investigation, based on a large corpus of stative possessives, shows that the now declining have got/got forms were much more prevalent in a Canadian context than had been previously attested. Literary evidence strongly suggests both have got and got were prominent in the speech of 19th-century Nova Scotian Loyalists, and evidence from sociolinguistics interviews and oral histories confirms that the forms were in fact the majority forms in the first decades of the 20th century.

The present study has also shown the consistency of the effect of subject type on variation between the stative possessive forms across Canadian varieties, with a specific constraint hierarchy not found in old world Englishes. In Cape Breton, as in the outlying Ontario communities, and all other varieties studied so far, subject reference is not a significant predictor of have or have got/got use. This perhaps indicates that the constraint was a specific development in Toronto.

Comparing the striking similarity between Cape Breton and Ontario’s rate and variable rules for have use diachronically with their difference in rate and variable rules diachronically for be like reveals a stark contrast between a transmitted and a diffused change. Cape Breton proceeds in lockstep with Ontario with respect to stative possessives and shows the same very specific subject type constraint. With be like on the other hand, Cape Breton is about a generation behind Ontario and shows no person or sex contrast. Chapters 8 and 9 will show a similar pattern. Anyone attempting to claim that these three changes in progress were diffused to Cape Breton from Inland Canada, like be like was, must then account for this strong parallelism.
Chapter 8

Deontic Modality

8.1 Introduction

The variation between *have* and *have got/got* forms is not limited to contexts in which the forms denote possession. Variation also occurs when the meaning is instead one of necessity, requirement, or obligation. As with stative possessives, the trajectories of change within the deontic modality system differ in the British Isles and North America, so it is particularly germane for assessing the genetic relationship between varieties of English.

8.2 The Variable

In Cape Breton English, deontic modality is usually expressed using one of four forms: *must*, *have to*, *have got to*, or *got to*. Although *need to* and periphrastic forms do occur, they occur at a much lower frequency than the four forms above.

(53)  
must  I think, any business, I don’t give a hell if it’s running a store, or a steel plant or whatever, it **must** make a profit. It cannot keep running at a loss. (Male born 1940, SW)

have to  I have to say, I am a slacker. (Female born 1993, PI)

have got to  People have got to learn to change, to start facing reality. (Male born 1955, SW)

got to  I don’t drive around that much, but of course at 90 years old I got to be careful. The very first accident that I get in now I’m through. Doesn’t make any difference or not if it’s my fault. (Male born 1921, PI)

need to  I’ll put that away in some filing cabinet for when I need to explain to my kids or something. (Female born 1987, PI)

other  What happens there is a welder goes through a course and he is required to take a test, a provincial test, and that is done every twelve months or upon the request of the supervisor who figures that his work is not up to standard. (Male born 1947, SW)

As with stative possessives in the previous chapter, the phonetically reduced forms ’*s got to*/’ve got to* and *gotta* take part in the variation.¹

(54)  
have got to  “Only thing,” my wife told me, ”you’ve got to give the foreman a bottle of rum.” (Male born 1910, ST)

have got to  He said, ”That’s fine. He’s gotta pay me too, or I’m gonna squeal on him.” (Male born 1907, ST)

got to  So when you go, you know you gotta, you gotta move the railcars around, stuff like that, yeah, sure. (Male born 1952, PI)

Again, as with stative possessives, the bare *got* variant appears in the local folk dictionary *Da Mudder Tongue* as an example of a Cape Breton idiom or vocabulary item. Not only does *got to* have its own definition, but it

¹. Phonetically reduced forms without *got*, e.g., ’*s to*/’ve to do not occur in the data.
is used quite frequently in the example sentences for other terms. Unlike stative possessives, however, the have variant also has an entry.2

(55) GODDA — A compulsion or having to do something, “I just godda get out fishin soon bye or I’ll have a connipation fit.” (Gray 2007, 24)
DON’T GODDA — Not obliged to do. “I don’t godda go to work Sun Dee.” (Gray 2007, 19)
HAFTA — You are required to do this. “I hafta clean up the gradge before I can go fishin.” (Gray 2007, 26)

There is a separate entry for have (56) — which is clearly intended to emphasize its pronunciation not meaning — that includes an example where have expresses deontic modality, but does not include an example where have expresses a stative possessive meaning.

(56) HAF — Another utility word with a double Caper meaning. Fifty percent of a part or having to do. “I had to give her haf me poker pot winnins.” The second example would be, “I can’t go fishin, I haf ta modda lon.” (Gray 2007, 26)

There is a clear linking of these phonetically-reduced forms of have, have to, got to, etc. to a local identity, and positions that local identity in opposition to the mainstream (Canadian or otherwise). The use of pronunciation spelling reinforces the association of these forms, and local speech more generally, with informality and a low education level, i.e., non-standardness. Da Mudder Tongue and Anudder Mudder are intended to be comedic,3 so there is also the clear association between local speech and the positive qualities generally associated with local affinity, group solidarity and/or covert prestige (Trudgill 1972). The inclusion of both the got and have forms with deontic meaning, but only the got form with stative possessive meaning; however, may indicate that there is a different local evaluation of the variants across these two variables. The distribution of the have (to) and got (to) variants with deontic and possessive meanings may reflect this potentially-differing evaluation.

### 8.3 Historical Perspective

The oldest of the modern forms for encoding deontic modality is the modal verb must. Beginning in the Late Middle English period the semi-modal construction have to began to compete with must for expressing the same meaning (see Crowell 1955, 69; Krug 1998; 2000; Traugott 1999). Both forms also encode epistemic modality, and the robust grammaticalization of have to as a semi-modal may coincide with must’s expansion into this meaning in the Early Modern Period.4

Have got to and got to enter the deontic modality system in the 19th century (F. T. Visser 1963–1973, 479) — mirroring, though slightly later than, the expansion of have got and got for stative possession.

The OED labels the use of bare possessive got and deontic got to as “colloquial,” “regional,” and “non-standard” (Oxford English Dictionary 2016). The very first attestation of bare got to, (57), in the OED actually comes from Nova Scotia author Thomas Chandler Haliburton (see Section 4.1).

(57) “All they got to do is, to up Hudson like a shot…and home in a liner, and write a book.” Haliburton 1836–1838, 2nd Series – V, ‘Travelling in America’

But Haliburton did not just employ bare got to for expressing deontic modality. His characters showed the full suite of deontic modality variants.

---

2. There is also an entry for must: MUSTA meaning ‘must have’, however this use of must is epistemic not deontic.

3. From the introduction to Da Mudder Tongue: “The idea for this book originated from another project of mine ‘The Official Cape Breton Joke Book’ which I published a few years ago. As I gathered the jokes from around the Island, I thought the phrases and words were sometimes as funny as the jokes” (Gray 2007, 1).

4. The timing and stages of development of semi-modal have to are not agreed upon among researchers (cf. Plank 1984, 320; Mitchell 1985, 950–53; Traugott 1999, 8; Krug 2000, 54; Brinton 1991, 11-12); however, have to’s status as a semi-modal encoding deontic modality is unambiguous by the 19th-century transporting of English to Cape Breton.
a. This Quaker, a pretty knowin’ old shaver, had a cause down to Rhode Island; so he went to Daniel to hire him to go down and plead his case for him; so says he, “Lawyer Webster what’s your fee?” “Why”, says Daniel, “let me see, I have to go down south to Washington, to plead the great Insurance case of the Hartford Company—and I’ve got to be at Cincinnati to attend the Convention, and I don’t see how I can go to Rhode Island without great loss and great fatigue.” (Haliburton 1836–1838, 1st Series — No. IV, ‘Conversations at the River Philip’)

b. When a man has to run ever so far as fast as he can clip, he has to stop and take breath; you must do that or choke. So it is with a horse; run him a mile and his flanks will heave like a blacksmith’s bellows: you must slack up the rein and give him a little wind, or he’ll fall right down with you. (Haliburton 1836–1838, 1st Series — No. XXIII, ‘The Blowin’ Time’)

The variation between must, have to, have got to and got to that occurs in the Sam Slick stories does not occur in the works of Thomas McCulloch (1821–1822/1990). McCulloch only uses must for expressing deontic modality — even though he seems to use stative possessive have got/got stylistically with Loyalist characters.

Perhaps the lack of have got to and stigmatized got to in Scottish-born McCulloch’s works is unsurprising given that these forms are only now making inroads into Scotland and Northern Ireland (Tagliamonte and Smith 2006, see below; Tagliamonte 2013a), and that the specific Nova Scotians McCulloch satirizes, though fictional, are generally believed to represent Scottish-settled Pictou County (New 2003, 61). But McCulloch’s similar non-use of have to, considering it had already been in use in the English language for well over a century and is the majority variant in conservative Scottish English, is perhaps more surprising. Contemporary analyses of the speech of older Scottish speakers from the Buckie corpus (Tagliamonte and Smith 2006; Tagliamonte 2013a) show no use of must for deontic modality and over 80 percent use of have to.

(59) A tavern must be open at all hours, and to all kinds of company. Irregularity in eating and sleeping requires the comfort of drinking. (McCulloch 1821–1822/1990, Letter 3, (as written by Mephisbosheth Stepsure))

Corpus-based research on a small sample of pre-Confederation Canadian literature from Inland Canada reported that the main variation was between must and have to, with very little have got to or got to (Dollinger 2008). This study, however, included both epistemic and deontic contexts (as will be discussed below must is used nearly categorically in epistemic contexts), and therefore the preponderance of must relative to have to is inflated.

8.3.1 The McKinnon Texts

Like The Letters of Mephisbosheth Stepsure the McKinnon Texts overwhelmingly employ must for expressing deontic modality. There is only one instance across the McKinnon Texts of deontic modality that is not must. It is from the rural Scottish Backlander character named Neil. The form is be have to or have have to, as in (60). This is not exactly have to, or have got to/got to, but it is a form using have for deontic modality that is not must. As with other variables, McKinnon does not always employ forms 19th-century Scottish English speakers are known to have used. This is apparent when the character Neil, for example, uses possessive have got, as in (42) in Chapter 7. Neil also uses the innovative will future, and uses it in first person, considered prescriptively incorrect in the mid-1850s and contra the overall pattern of future temporal reference expressions in the McKinnon Texts more generally. Neil’s use of be have to or have have to to likely results from McKinnon being to some degree aware of deontic modality have forms, associating them with non-standard speech, and employing them (albeit incorrectly) in the speech of his low status backlander character.

(60) “Al pe sure enough, and more sure, too,” replied Neil, “he’s have to clos on fat she’ll say, and he was pretty mans too.” (McKinnon 1852, 63)

Only three of the remaining 82 instances of a character expressing deontic modality — all must — are contexts in which the obligation is objective or external to the speaker. This is likely because 76 percent ($N = 62$) of the
tokens occur with first person subjects and subjective, or personal, obligation are often correlated.

### 8.4 Synchronic Perspective

Variation in the deontic modality system has been studied on both sides of the Atlantic, and the emergence, rise, and spread of each variant has been documented in both North American and British varieties.

Krug (1998; 2000), for example, finds, based on an analysis of several corpora of written (and to a lesser degree spoken) British English, that *have got to* and *got to*° are on the rise in Britain. This finding was confirmed by Jankowski’s (2004) analysis of 20th-century British plays. In the plays, *must* is used over 60 percent of the time prior to the 1950s and decreases to under 20 percent after the late 1970s, with both *have got to* and *got to* concomitantly increasing in parallel. Conversely, in 20th-century American plays there is robust variation between *must* and *have got to* until the mid 1920s, at which point *must* begins rapidly declining relative to all other forms. From the mid 1920s to the early 1950s *have got to* is the dominant variant, but after this time *have to* takes over, reaching about 60 percent usage by the late 1970s. After the 1950s *got to* also takes over as the main secondary variant.

Photograph 7: *Must* for expressing deontic modality in signage, Canadian English. (Photographed at the Voi-sey’s Brook Dog Park, Portugal Cove-St. Philip’s, Newfoundland, by the author, August 18, 2016).

Photograph 8: *Must* for expressing deontic modality in signage, American English. (“You MUST push button to call for walk signal” photographed at the intersection of New York Avenue and North Capitol Street NE, New York, USA, by sglazerman, August 18, 2016). CC Image courtesy of sglazerman on Flickr https://flic.kr/p/ftFVXX.

Tagliamonte and Smith’s (2006) analysis of a cross-section of speakers from England, Scotland and Northern Ireland find an overall decline in deontic *must* in apparent time. The few communities in the study where deontic *must* is used to any degree, outside of formulaic expressions like *I must say*..., are Henfield and York (see Map 6 on p. xviii). The authors speculate that this may be due to these communities’ relatively-larger percentage

---

5. Or perhaps more specifically *have gotta* and *gotta*
of middle-class residents, and the association of deontic *must* with formality. Across communities there is robust variation between *have to* and *have got to*; however, in the most northern communities, Portavogie and Culleybackey in Northern Ireland and Buckie in Scotland, the *have got* variant is only nascent, used only among younger speakers. Rather than finding evidence of an ongoing change from *have to* to *have got to* to *got to* (i.e., *gotta*, as predicted by Krug 1998; 2000), Tagliamonte and Smith (2006) find evidence of stability or a potential diachronic pivot back to *have to*.

Of particular relevance to the overriding question of this dissertation is the lack of *have got to/got to* forms among the conservative Scottish and Irish speakers from Buckie and Culleybackey/Portavogie. As with stative possessives, if the *have got to/got to* forms exist within the Cape Breton data, and these forms are not present in speech of speakers from conservative Scottish and Irish varieties, than Cape Breton must have inherited these forms from a different input source, acquired them through diffusion, or developed them independently as part of linguistic drift.

There have been several synchronic studies of deontic modality in both Inland Canada and Atlantic Canada. Tagliamonte (2000), for example, analyzes variation in the deontic modality system among speakers in the corpus of African–Nova Scotian English (Poplack and Tagliamonte 1991). Among African–Nova Scotians living in North Preston (outside Halifax) and Guysborough Enclave (in northern mainland Nova Scotia, see Map 4 on p. xvi) bare *got to* is used about twice as often (64–66 percent) as the second most common variant *have to* (32–34 percent). Both *have got to* and *must* are infrequent, but both are present in the data (see Figure 8.2). And while in both communities *must* is used less than five percent of the time in deontic modality contexts, it is used over 89 percent of the time in epistemic modality contexts.

Tagliamonte and D’Arcy (2007b) track changes in the deontic modality system of 76 speakers from the *Toronto English Archive*. The authors find a monotonic increase in *have to* use over apparent time, growing

---

6. In many English-speaking communities deontic *must* remains the pragmatically-appropriate variant on signage, *e.g.*, *All dogs must be on lead*. See Photographs 7–9.

7. Data from these cities was combined in Tagliamonte and Smith’s (2006) study.

8. Tagliamonte (2000) analyzes *got to* and *gotta* as separate variants. Here they are combined.
from 63 percent among speakers born prior to 1945 to 90 percent among speakers born after 1987. Only the oldest speakers in the data (born before 1945) use must to express deontic modality, and for all speaker age groups (born <1945, 1946–1973, 1974–1987, >1987) the main secondary variant is bare got to, which is used at a higher rate than either have got to or need to.

Among the three outlying Southern Ontario communities examined by Tagliamonte and Denis (2014) (see Map 5 on p. xvii) have to is also the dominant form, used between 60 and 80 percent of the time across communities. The authors paint a nuanced picture, however, of the deontic modality system in the three communities based on their statistical analyses. For example there is still some, albeit scant, must usage in Belleville, which the authors attribute to the community’s conservatism and relatively higher rate of Loyalist settlement at found-
ing (Tagliamonte and Denis 2014, 110; cf. Dollinger 2008). Belleville is also the only community where have got to exceeds 10 percent usage. Burnt River and Lakefield, both with stronger Ulster Scots roots, have higher — and in the case of Lakefield, stable — rates of have to relative to Belleville, along with a paucity of have got to and no must (Tagliamonte and Denis 2014, 110; cf. Tagliamonte and Smith 2006). The authors further conclude that the three communities are changing in parallel with Toronto towards increasing have to use, and given that the underlying grammar of the deontic modality system is consistent across communities, this parallel change results from transmission from a common predecessor variety (Tagliamonte and Denis 2014, 110).

In their analysis of expressions of deontic modality among Toronto speakers of various ethnic backgrounds, Walker and Hoffman (2016) find no change in have to use between older and younger speakers of British/Irish ethnic origin. Both groups use have to about 60 percent of the time. The authors did find, however, an increase between these two groups in the use of need to, which grows from about 10 percent use among older speakers (>40 c. 2010) to about 20 percent among younger speakers (18–30 c. 2010). The rates of use of each variant in the deontic modality system were variable among speakers of various ethnic backgrounds; however, have to is the dominant form across groups (59 percent overall), followed by need to (20 percent), got to (13 percent), must (5 percent) and have got to (3 percent). Of note is that among these speakers, as with speakers in Toronto, Burnt River and Lakefield (but not Belleville), got to is used more frequently than have got to.

The synchronic and historic picture of the deontic modality system in English leads to several implications with direct relevance to assessing the genetic relationship between Cape Breton English and Inland Canadian English or Scottish/Irish/British English. There is strong evidence that Inland Canadian English inherited the have got to/ got to forms from late-18th/early-19th century Loyalist immigration. In communities with relatively higher rates of Loyalist immigration, like Belleville, these forms may have been more firmly entrenched in the community grammar and so these communities show a greater frequency in their vestigial use of the forms. Communities with a relatively larger amount of 19th-century Scottish/Irish immigration will likely have a lower rate of have got to/got to because the forms are only beginning to penetrate conservative Scottish and Irish English varieties. Furthermore, the use of must in conservative Scottish/Irish varieties is nearly non-existent in deontic contexts, so communities with relatively greater Scottish/Irish immigration may also show low rates of this variant.

8.5 Methods

As with the previous chapter the methods for extracting, coding and analyzing deontic modality tokens from the Storyteller, Steelworker and Post-Industrial corpora were intentionally modelled on those employed by previous studies, e.g., Tagliamonte and Smith (2006), Tagliamonte and D’Arcy (2007b), Tagliamonte (2012b), and Tagliamonte and Denis (2014), for maximum comparability.

All instances of must, have, have got, got and need to (including phonetically reduced forms) with a deontic modality meaning in the present tense were extracted from the data. These verbs in negative and interrogative phrases were extracted for qualitative analysis, but were excluded from the statistical analysis below, as the above studies do not include them in their analyses. Likewise, these verbs in epistemic modality contexts were extracted for comparative purposes, but are not included in the statistical analyses.

Past and future morphological contexts were not extracted because they were categorically have to or another periphrastic construction like was/will be required to. Must and have got to/got to cannot felicitously occur with

---

9. The authors do not comment on the rates of use of got to, have go to or must among individual groups, instead only listing rates of have to and need to among each group. The high overall rate of need to results from particularly high rates of this variant among speakers of Greek (about 20 percent), Portuguese (about 30 percent) and Punjabi (about 40 percent) ethnic origin.
a deontic meaning with anything other than present tense morphology (as in (61)), so morphologically future and past contexts are outside of the envelope of variation. The use of the present tense with a future or past temporal referential meaning as indicated by temporal adverbs, on the other hand (e.g., *I must/have to/have got to/got to leave tomorrow*), were not excluded in the analysis.

(61)  
Past *I must/must-ed leave yesterday.*  
Past *I had got to leave yesterday.*  
Past *I got/got-ed to leave yesterday.*  
Future *I will must leave tomorrow.*  
Future *I will have got to leave tomorrow.*  
Future *I will got to leave tomorrow.*

Also excluded were the non-deontic expression have/have to/got to do with… (see Jespersen 1961a, 207) and got to where the meaning was either ambiguous or non-deontic:

(62)  
a. My mother does the crossword every day. But she’ll ask me, “What’s the word?”, you know, and usually, again, it’s something to do with something modern like that. It’s a word that’s—, you know, they’ll ask something that has to do with a computer background. (Female born 1961, PI)
b. The whole thing was repetitious. You got to be very blasé about the whole thing. (Male born 1916, SW)
c. They’d never let you use it. I never got to use it. (Male born 1927, SW)

The variants preceded by other modals were not included as Cape Breton English does not allow double modal constructions (e.g., *could must*; *might must*), even though it does allow a modal plus a semi-modal (though with different meaning). This is another reason why modal future temporal reference expressions like *shall* and *will* are excluded:

(63)  
a. They should have to work right from the bottom to the top. (Male born 1943, SW)  
b. A lot of people lost their legs and arms and was really hurt because you’d have to climb under the cars, open the doors. (Male born 1927, SW)

8.6 Results

A total of 2,498 tokens were extracted from the four Cape Breton English corpora. There were 83 tokens of stative possessive verbs in the dialogue in the McKinnon Texts (discussed above). The 2,498 tokens includes 203 tokens in either negative, interrogative, or metadiscursive contexts (discussed below), which, following a qualitative analysis, were excluded. The remaining 2,212 tokens from the Storyteller, Steelworker, and Post-Industrial corpora were used for regression analysis.

8.6.1 Deontic Modality and Discourse

Following the methods used in the studies listed in Section 8.5, tokens that occur in metadiscursive contexts were excluded from analysis. In the studies listed above tokens like *I must say…* and *I must admit…* were excluded; the authors label them as formulaic, invariable constructions (e.g., Jankowski 2004, 92). Here, instead, *I must say…*, *I must admit*, and similar constructions were labelled as metadiscursive and laid aside for qualitative analysis. The criteria for identifying the 86 metadiscursive tokens in the data were that they must appear on the left edge of a clause or as a tag, and that they must provide information about the role of the basic message of the following (or preceding) utterance as a discourse activity. Fraser (1997) labels these types of expressions “perlocutionary markers.” These markers are external to the primary discourse and are instead a meta-discursive pragmatic commentary. For example, compare the use of *have to say* in the two examples below:
Primary Discourse  You’ve got a degree of people that don’t want to do their job and you’ve got a degree of people that you never have to say anything to them. The ‘80-20 rule’, you know, where you spend eighty percent of your time going after the twenty percent that don’t do anything, and you only give twenty percent of your time to the eighty percent that are doing their job. (Male born 1957, SW)

Metadiscourse  They had ladies down there in the steel plant operating cranes. They were very efficient. They never missed time. They never came on the job half-impaired like some of the men did. I have to say this; this is the truth. They worked with me in the brickshed, wheeling wheelbarrows and bricking, and they worked just as hard as anyone else. So I can’t say nothing about women. (Man born 1927, SW)

In the first example the speaker is describing a group of people he works with to whom a manager would never be required to provide instructions. The deontic modality variant have to occurs as part of the basic meaning of the utterance. In the second example the speaker is introducing a declarative statement about female steel plant workers by first indicating to the interviewer that he feels compelled to make the declaration. Table 8.1 lists the expressions in the data that meet the two criteria for being a perlocutionary marker.

Table 8.1: Distribution of deontic modality verbs in Cape Breton English — metadiscursive contexts by frame.

<table>
<thead>
<tr>
<th>Expression</th>
<th>must</th>
<th>have to</th>
<th>have got to</th>
<th>got to</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I ______ say,…</td>
<td>42</td>
<td>11</td>
<td>1</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>I ______ admit,…</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>I ______ tell you,…</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>I ______ be honest,…</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>I ______ mention,…</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>I ______ ask,…</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>You ______ remember,…</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>You ______ understand,</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>You ______ realize,…</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>You ______ recall,…</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>24</td>
<td>3</td>
<td>5</td>
<td>86</td>
</tr>
</tbody>
</table>

Table 8.2 lists the number of tokens of each deontic modality variant in metadiscursive contexts by speaker decade of birth. Among speakers born in the 1940s and earlier these contexts occur nearly exclusively with must, and among those born in the 1960s and afterwards, nearly exclusively with have to.

Table 8.2: Distribution of deontic modality verbs in Cape Breton English — metadiscursive contexts by decade of birth.

<table>
<thead>
<tr>
<th>Decade of Birth</th>
<th>must</th>
<th>have to</th>
<th>have got to</th>
<th>need to</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1890s</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1920s</td>
<td>19</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>1930s</td>
<td>14</td>
<td>2</td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>1940s</td>
<td>17</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>1950s</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>1960s</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>1970s</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>1990s</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>27</td>
<td>3</td>
<td>5</td>
<td>90</td>
</tr>
</tbody>
</table>

While these metadiscursive commentary markers show variation between deontic modality variants, they
were not included in the statistical analysis below because of their disproportionate use by older speakers with
must in much of the data. By excluding metadiscursive contexts, the specific envelope of variation considered
for statistical analysis is primary discourse. Previous studies do not label their envelope of variation as such, but
given these studies’ exclusion of I must say…, etc., their envelope of variation is close to exclusively primary
discourse anyway. Limiting the present analysis to primary discourse, rather than identifying I must say… and
I must admit… as ad hoc exclusions, is instead linguistically motivated.

Across studies of Canadian, American and British English (including Scottish and Irish English) these metadis-
cursive contexts have been excluded because of must’s complete or nearly complete dominance within them
(e.g., Jankowski 2004, 92; Trousdale 2003, 276–277; Tagliamonte 2004, 38; Tagliamonte and Smith 2006; Taglia-
monte and D’Arcy 2007b, 73). However, the fact that this dominance has been found across varieties suggests,
as Tagliamonte and D’Arcy (2007b) and others point out, that must has retreated in English generally to highly
circumscribed contexts.

The preponderance of must in these formulaic metadiscursive contexts likely results from a common phe-
nomenon of obsolescing features: that they often “get left behind, fixed in formulaic utterances or certain dis-
course rituals”10 (Tagliamonte and D’Arcy 2007b, 73; see also Tagliamonte 2004, 50). Perlocutionary markers
convey pragmatic rather than denotational meaning, and so are likely produced and processed with less focus
on their compositional syntactic/semantic meaning. This is what makes them “formulaic” and for this reason
innovations may take longer to (or may never) penetrate them.

But penetration is not impossible, so metadiscursive contexts must still be taken into account when ex-
amining changes to the deontic modality system — even if they are excluded from statistical analyses. In both
the Cape Breton data and in the results reported by Tagliamonte and D’Arcy (2007b), among speakers born in
the second half of the twentieth century, have to takes over as the majority form in metadiscursive contexts,
signalling must’s continued obsolescence for expressing deontic modality in spoken English, even in this last
hold-out context. That peripheral Cape Breton and innovative Toronto exhibit the change from must to have
to in metadiscursive contexts at the same time (rather than showing evidence of geographic diffusion) suggests
that this change was inevitable as part of the — as will be argued below — parallel transmission in both com-
munities. A future independent cross-varietal study of changes to deontic modality in metadiscursive contexts
(including Canada, Britain, United States, etc.) may prove fruitful in the study of linguistic obsolescence.

8.6.2 Deontic and Epistemic Modality

In contrast to the similarity between Toronto and Cape Breton English with respect to metadiscursive contexts,
in epistemic modality11 contexts Cape Breton does not match Toronto.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>must</th>
<th>have</th>
<th>have got</th>
<th>got</th>
<th>need to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Steelworker</td>
<td>201</td>
<td>92</td>
<td>10</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>41</td>
<td>93</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>242</td>
<td>92</td>
<td>11</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

10. Their “terminal refugia” as in Footnote 4.
11. Epistemic modality is the area of mood concerned with knowledge and belief, including at least the expression of possibility, probability
and certainty (Trask 1993, 92). In contrast, deontic modality is the area of mood concerned with permission, obligation and necessity (Trask
1993, 76).
As Table 8.3 shows epistemic must occurs about 92 percent of the time relative to the other variants in the Steelworker and Post-Industrial corpora. The Storyteller corpus, which is the smallest of the three spoken corpora, did not contain any instances of epistemic modality, so it is not included in Table 8.3. The rate of must in epistemic contexts in Cape Breton is nearly identical to the rate among African–Nova Scotian speakers in North Preston (94 percent) and Guysborough Enclave (90 percent) reported by Tagliamonte (2000, see also Tagliamonte and D’Arcy 2007b, 69). The scant use of have to, have got to, and got to sets Cape Breton apart from both Toronto, where these forms are used about 45 percent of the time (Tagliamonte and D’Arcy 2007b, 69), and conservative Scottish and Northern Irish varieties, where they are never used at all (Tagliamonte and Smith 2006).

Epistemic Modality I don’t know what area we’re looking at here. I can see the coal piers or some kind of piers in the background. It must be an old picture. I don’t know if those trestles or piers are still there. (Male born 1947, SW)

Epistemic Modality There has to be a way of getting him back on. (Male born 1927, SW)

Epistemic Modality Pretty well anybody with a trade in there has got to have a little knowledge of it. (Male born 1944, SW)

Epistemic Modality [He wears] Lifts. It gotta be! (Female born 1987, SW)

Tagliamonte and D’Arcy (2007b, 70) point out that epistemic meanings arise from earlier deontic meanings and thus Toronto’s higher use of epistemic have to (18 percent) and have got to (18 percent) may represent a more advanced stage of these forms’ grammaticalization (see Heine and Kuteva 2002; Heine, Claudi, and Hünnemeyer 1991; Traugott 1989). In fact, of the multiple corpora, both spoken and written, surveyed by Tagliamonte and D’Arcy (2007b), have to and have got to are used drastically more frequently in epistemic contexts in Toronto. Increased grammaticalization of have to/have got to into epistemic contexts is likely indicative of the difference between innovative Toronto and peripheral Cape Breton — a difference that is not immediately apparent from the variation between these forms in deontic contexts.

8.6.3 Deontic modality verbs and do-support

While the presence of do-support with stative possessive have is a site of difference between North American and British English varieties, the use of do-support with deontic have to is the norm on both sides of the Atlantic (e.g., Quirk et al. 1985, 145).

Table 8.4: Distribution of deontic modality verbs in Cape Breton English — negative present-tense contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>don’t have to</th>
<th>haven’t got to</th>
<th>don’t need to</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steelworker</td>
<td>75</td>
<td>1</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>14</td>
<td>3</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>1</td>
<td>3</td>
<td>93</td>
</tr>
</tbody>
</table>

In negative contexts in the Steelworker and Post-Industrial corpora have to was used 96 percent of the time and do-support was categorical. There were no negative deontic modality tokens in the smaller Storyteller corpus. The one token of negative have got to occurred as haven’t got to, where do-support is not expected anyway. The 93 negative constructions were excluded from statistical analysis due to their near categoricity.

Interrogative tokens were also excluded from further analysis because they occurred nearly categorically with have to. The six “Other” tokens listed in Table 8.5 were interrogative statements constructed with either utterance final tags or intonation alone. Of those six, five are have to and one is need to.

12. There were also no epistemic modality tokens in the McKinnon Texts.
Table 8.5: Distribution of deontic modality verbs in Cape Breton English — interrogative present-tense contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>do I have to?</th>
<th>do I need to?</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storyteller</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Steelworker</td>
<td>8</td>
<td>3</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>1</td>
<td>6</td>
<td>21</td>
</tr>
</tbody>
</table>

8.6.4 Trajectories of Change

Table 8.6 reports the frequency of use of each of the five deontic modality variants in the Storyteller, Steelworker, and Post-Industrial corpora in non-interrogative present tense contexts with affirmative polarity in primary discourse. This data is also charted against decade of birth in Figure 8.3. What is immediately apparent is that among the oldest speakers, and in the oldest corpora, variation between forms is robust. As with stative possessives there is a high rate of have got to (dashed orange line with triangles) among those born prior to 1920. Also as with stative possessives, there is an increase in the have form, here have to (solid blue line with squares), over both real and apparent time. Finally, again as with stative possession, among speakers born in the second half of the 20th-century, the bare got form, here got to (dashed green line with circles), takes over as the main secondary variant. Figure 8.4 tracks moribund must (solid pink line with squares) and tertiary need to (dashed green line with triangles) across corpora and decade. While Walker and Hoffman (2016) suggest the rise of need to is a burgeoning change in progress (an observation not supported by Tagliamonte and D’Arcy 2007b), the Cape Breton data do not show this change. In fact, need to is essentially a non-entity in the Cape Breton deontic modality system, representing just 1 percent of the data (less than in any other Canadian data). Must is also only used about 2 percent of the time — a stark change to the near categoricity of deontic must in the McKinnon Texts.

The trajectories of change tracked in Table 8.6 and Figure 8.3 show that Cape Breton English is more closely aligned with the Inland Canadian data reported by Tagliamonte and D’Arcy (2007b), Tagliamonte and Denis (2014), and Walker and Hoffman (2016), rather than the data reported by Tagliamonte and Smith (2006) and Tagliamonte (2013a) for Scotland and Northern Ireland. Figure 8.5 collapses the three corpora and charts the rise of have to in Toronto (Tagliamonte and D’Arcy 2007b) and Buckie (Tagliamonte and Smith 2006) for comparison. The change in have to use in both cities is portrayed as step-wise because in those studies frequencies are provided for speakers grouped in age ranges that span multiple decades. The Cape Breton data has similar rates of have to use as Toronto speakers across apparent time. The Cape Breton data does not, however, match the frequency of have to use among Buckie speakers over apparent time. The oldest speakers in both Cape Breton and Toronto use the most have got to, while the oldest speakers in Buckie use the most have to.

The obvious outlier decade in the Cape Breton data is the 1980s. Speakers born during this decade use bare got to about 33 percent of the time, a rate much higher than any age group in any other Canadian community in the literature. This group of outliers will be discussed below in Section 8.7.2.

13. For Figure 8.3 and several other figures in this section, must and need to are not represented; however, their presence within the variable deontic modality system was included in the calculations used to create the figures.
Table 8.6: Distribution of deontic modality verbs in Cape Breton English — present-tense affirmative declarative contexts.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>must</th>
<th>have</th>
<th>have got</th>
<th>got</th>
<th>need to</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Storyteller</td>
<td>9</td>
<td>19</td>
<td>33</td>
<td>70</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Steelworker</td>
<td>46</td>
<td>3</td>
<td>1,287</td>
<td>69</td>
<td>339</td>
<td>18</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>4</td>
<td>2</td>
<td>145</td>
<td>63</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>2</td>
<td>1,441</td>
<td>68</td>
<td>379</td>
<td>18</td>
</tr>
</tbody>
</table>

Figure 8.3: Distribution of deontic modality verbs in Cape Breton English — variants by corpus and decade of birth.

The distribution of deontic modality forms provides yet more strong evidence that Cape Breton English can trace its roots to Loyalist English rather than (or in addition to) Scottish or Northern Irish varieties. While the *have got to/got to* forms are only incipient in conservative speech in the northern British Isles, in the current study, speakers born before the 1920s — the oldest speakers in the Cape Breton data, and the speakers who ought to pattern most like conservative Scottish/Irish speakers if Cape Breton English derives from those varieties — use *have got to* as their majority form. Tagliamonte and Denis (2014) contrast Burnt River/Lakefield with Belleville, explaining that the relatively higher proportion of Ulster Scots founders in the former and the relatively higher proportion of Loyalist founders in the latter, lead to differences in rates in *have to* between communities.

In the combined Cape Breton data, and for each of the three corpora (*Steelworker, Steelworker, Post-Industrial*), Cape Breton has a higher rate of Loyalist-origin *have got to/got to* forms than prototypically-Loyalist Belleville. The only reason Belleville has a lower rate of *have to* compared to Cape Breton is the abundance of *need to* in Belleville.

Strong additional evidence that the deontic modality system in Cape Breton derives from Loyalist rather than Scottish origins will come from a careful analysis of the grammatical constraints governing the use of each variant in the Cape Breton data.
Figure 8.4: Secondary deontic modality verbs across corpora of Cape Breton English — variants by corpus and decade of birth.

Figure 8.5: Distribution of stative possessive *have* in Cape Breton, Toronto and Buckie Variants by decade of birth. Comparative rates of *have to* in apparent time for Toronto (2004) and Buckie (1997), as presented by Tagliamonte and D’Arcy (2007b, Figure 4, Table 10) and Tagliamonte and Smith (2006, Table 10)
8.6.5 Linguistic Conditioning

As with the rise of stative possessive *have got/got*, the rise of *have got to/got to* in English is well studied, with a number of quantitative analyses of the variable in both North American and British varieties. These studies have isolated specific grammatical contexts in which the *have got/got* forms are favoured, and in which contexts traditional *must* and *have* are favoured. Again, as with stative possessives, the diachronic trajectories of each variant, as well as the specific favourable contexts differ in the two macro-varieties.

8.6.5.1 Subject Type

![Distribution of deontic modality verbs by subject in Cape Breton English](image)

Figure 8.6: Distribution of deontic modality verbs in Cape Breton English — variants by subject type.

In the literature on stative possessives, there is an hypothesis that *have got* arises historically when *have/has* begin to contract to [v] and [s] in pronominal contexts as a way to reinforce the relationship between subject and verb (Crowell 1955, 283). This initial use with pronominal subjects leads to a diachronic subject type effect, whereby at any given time the *have got/got* forms are more likely with pronominal subjects. This pattern has been attested in some communities, while more complex patterns have been found in others (see Section 7.5.3.1). In the Cape Breton data, for example, there is a subject effect, but the relevant distinction is between noun phrases/third person plural pronouns, first and second person singular pronouns, and third person pronouns. The extent to which this pattern also exists for deontic modality is testable.

(67)

Noun Phrase Far as I’m concerned, **whatever goes into that pit** has to come out the same temperature, cold or hot. (Male born 1946)

3rd Person Plural Pronoun People leave here from Cape Breton and go to Halifax, lots of work up and around there, but not good paying jobs. They are up there for awhile and **they** have to come back. They can’t afford to live up there and that’s a proven fact. (Male born 1932)

1st Person Pronoun So he would call me, “You having problems?” If I am having problems, “Yes, **I** have to come off. Okay?” (Male born 1930)
2nd Person Pronoun  Now, you, you have to come a little closer because I have a problem hearing you. (Male born 1923)
3rd Person Singular Pronoun  Tonight I’ll go out and they’ll say such and such is broke down. So you’ll go and you’ll check all the electrical stuff and it’s still not working. Then I got to pick up the phone and call the technologist and he got to come over and get into the computer and see the problem. (Male born 1954)

Figure 8.6 shows the proportion of each variant by subject type, with third person singular and third person plural pronominal subjects separated following the patterning for stative possessives. At first blush, the deontic modality system is unlike the stative possessive system in that third person singular pronominal subjects, which show the least amount of have, are more advanced towards have to relative to first and second person pronominal subjects. That being said, the stative possessive and deontic modality systems are similar insofar as for both changes NP subjects and third person plural pronominal subjects pattern similarly and show the greatest rate of have.

8.6.5.2 Subject Reference

Figure 8.7: Distribution of deontic modality verbs in Cape Breton English — variants by subject reference.

Research on changes in the deontic modality system has also found a significant effect of subject reference on variation between have to and have got to/got to; however, the direction of the effect is not consistent across studies. Jankowski (2004, 101) find that in both American and British plays have to is highly favoured by generic subjects. Tagliamonte and D’Arcy (2007b) find generic subjects favour have to in their Toronto data, but the effect was only operational among speakers aged 30 to 59,14 while Walker and Hoffman (2016) find generic subjects disfavour have to in their corpus of Toronto speakers from different ethnic backgrounds. Tagliamonte and Smith (2006) analysis of British varieties also finds that generic subjects disfavour have to relative to other subjects.

---

14. The patterning of generic versus other subjects is slightly more complex than generic vs. specific vs. existential. This complexity will be addressed below.
Table 8.7 shows the rates of each variant by subject reference. In Cape Breton there is a slightly higher rate of *have to* with generic subjects than with specific subjects, but both have a lower rate than existential subjects. This is perhaps to be expected, as existential subjects are all third person pronominal subjects (and all singular in the data) and all generic subjects are second personal pronominal subjects. As shown in Figure 8.6 these two subject types are the least likely to occur with *have to*, and most likely to occur with *have got/got to*. Subject reference is non-orthogonal with subject type and any statistical analysis that includes subject type and subject reference must take this into account.

Page 113

<table>
<thead>
<tr>
<th>Variant</th>
<th>Objective</th>
<th>Subjective</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>must</em></td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td><em>have to</em></td>
<td>71%</td>
<td>57%</td>
</tr>
<tr>
<td><em>have got to</em></td>
<td>18%</td>
<td>18%</td>
</tr>
<tr>
<td><em>got to</em></td>
<td>9%</td>
<td>2%</td>
</tr>
<tr>
<td><em>need to</em></td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>(Total N = 1649)</td>
<td></td>
<td>(Total N = 485)</td>
</tr>
</tbody>
</table>

Figure 8.8: Distribution of deontic modality verbs in Cape Breton English — variants by obligation type.

The final constraint on deontic modality variation that has been attested in past studies is obligation type. When the obligation or necessity expressed by the deontic modality form is externally-motivated (i.e., objective) *have to* is favoured relative to contexts where the obligation is internally-motivated (i.e., subjective). Tagliamonte and D’Arcy (2007b) and Tagliamonte and Denis (2014) find this pattern in Toronto and in the three outlying Southern Ontario communities, as do Tagliamonte and Smith (2006) and Tagliamonte (2013a) across the British Isles. Walker and Hoffman (2016) do not test obligation type.

(68) Generic Yeah, it’s a lot of work, like, the work it is not hard. It’s just *you* need to find the time and the effort to do it. (Female born 1991, PI)

Specific He said, “*You* have to understand something. This is my only livelihood. It’s the only thing I’ve done since I was sixteen years old. *I* have to go, if I can save the mine.” (Female born 1966, PI)

Existential Took us so many weeks. Came to the realization that a community, or, this community cannot exist primarily on basically tertiary industry, on service industry. We came to the conclusion in fact, *there* has to exist a primary industry. (Female born 1948, PI)

8.6.5.3 Obligation Type

(69) Objective They just take it from you and they take it to the office and then you *have* to get a parent to come and sign for it to take it back out. (Male born 1991, PI)
Subjective  And we were at his house. Anyway, it was like, probably like three in the morning and he's like "Ah, I got to piss." And somebody was in the washroom when we were at his house. So he pissed in the glass, right? And he just put it on the, like, put it on the table and walked away and was there sitting on the computer, [laughs] like sitting there for like an hour. And I'm just, we were just sitting there talking and stuff. And then he gets up and he comes over and he's like "Oh man, I'm thirsty." And he's like "Yeah, sweet. There's still some wine left." And he picked up his glass, and started drinking it.[laughs] (Male born 1983, P9)

This constraint too overlaps with subject type and subject reference. Deontic modality with generic subjects — which are always second person pronouns — is almost always objective. Likewise, most subjective obligation contexts occur with first person subjects. Therefore, to test the effects of these three non-orthogonal constraints, subject type, subject reference and obligation type were combined into a single parameter, which compares generic subjects (almost always objective), noun phrases with objective obligation, noun phrases with subjective obligation, pronouns with objective obligation and pronouns with subjective obligation. This method of dividing tokens is consistent with strategies employed by Tagliamonte and D’Arcy (2007b), Tagliamonte and Denis (2014), and Walker and Hoffman (2016). The distribution of variants across these five categories is displayed in Figure 8.9.15

Distribution of deontic modality verbs by subject and obligation type in Cape Breton English

Figure 8.9: Distribution of deontic modality verbs in Cape Breton English — variants by subject and obligation type.

8.6.5.4 Logistic Regression Analysis

To test the combined effects of subject/obligation type several statistical modelling methods were employed. Based on the distribution of variants apparent in Figure 8.3 the three corpora — Storyteller, Steelworker and Post-Industrial — were analyzed separately. This strategy was supported by a conditional inference recursive partitioning tree analysis (Section 5.1, p. 44), Figure 8.10, which firstly found a significant difference between the Storyteller corpus ($p < 0.001$), which contains significantly more non-have to tokens, and the other two

---

15. This configuration does not permit exploring whether the unique Canadian subject type constraint hierarchy is also operational for deontic modality. Additional coding and analysis — discussed in Section 8.7 — was therefore performed.
Corpora. The tree analysis further found a significant difference between the Post-Industrial and Steelworker corpora ($p = 0.033$) with the latter having a higher probability of *have to* than the former. Splitting the three corpora unfortunately left only 47 tokens from the Storyteller corpus and 232 tokens from the Post-Industrial corpus, which, as will be shown below, results in slightly different models for each corpus.

![Figure 8.10: Conditional inference recursive partitioning tree — *have to* vs. *must/have got to/got to* in Cape Breton English. Parameters included: corpus.](image)

A secondary conditional inference recursive partitioning tree analysis of the Steelworker and Post-Industrial corpora, Figure 8.11, found two shock points in apparent time. In the Steelworker corpus there is a significant difference between speakers born in or before 1935, who use less *have to*, and speakers born after 1935. There is also a difference between speakers born in or before and after 1948, with the younger speakers using more *have to*. Likewise, in the Post-Industrial corpus there is a significant difference between speakers born in or before 1990, who use less *have to*, and speakers born after 1990, who use more *have to*. Of note, the speakers born before 1990 in the Post-Industrial corpus include speakers born before 1935 and 1948, as well as speakers born in the 1980s, who appear to be using a lot more *got to*. This accounts for the lower rate of *have to* among pre-1990 speakers in the Post-Industrial corpus compared to the post-1935/post-1948 speakers in the Steelworker corpus.

Table 8.7 shows the results of an analysis of deviance on a mixed-effects logistic regression with subject/obligation type, year of birth, and sex as fixed effects\(^{16}\) and speaker as a random effect. This analysis confirms several facts about deontic modality variation in Cape Breton English. First, for the Steelworker and Post-Industrial corpora, subject/obligation type is not only significant but the predictor with the greatest magnitude of effect on the variation (as indicated by the $\chi^2$ value). Second, among both the Steelworker and Post-Industrial corpora, age is also a significant predictor of the variation, confirming the trends observed in the distributional data that the variation in the deontic modality system represents a change in progress (at least in apparent time). Lastly, sex is a significant predictor for the Steelworker data, but not for the Post-Industrial data. The distribution of tokens among men and women may have resulted in this finding, as will be discussed below.

Table 8.8 presents the results of three mixed-effects logistic regression analyses, one for each corpus. The

---

16. There were only three female deontic modality tokens from the Storyteller corpus, so this parameter was not included as a fixed effect for the Storyteller model.
Figure 8.11: Conditional inference recursive partitioning tree — *have to* vs. *must/have got to/got to* in Cape Breton English. *Steelworker* and *Post-Industrial* corpora. Parameters included: year of birth, sex.

Table 8.7: Analysis of deviance — Wald $\chi^2$ test for full model, data partitioned by corpus. *Have to* for expressing deontic modality in Cape Breton English.

<table>
<thead>
<tr>
<th>Parameter (Factor Group)</th>
<th>df</th>
<th>Storyteller $\chi^2$</th>
<th>p-value</th>
<th>Steelworker $\chi^2$</th>
<th>p-value</th>
<th>Post-Industrial $\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject/Obligation Type</td>
<td>2/4/3</td>
<td>1.10</td>
<td>$5.80 \times 10^{-01}$</td>
<td>17.20</td>
<td>$1.80 \times 10^{-03}$</td>
<td><strong>11.37</strong></td>
<td>$2.10 \times 10^{-04}$</td>
</tr>
<tr>
<td>Year of Birth, centred</td>
<td>1</td>
<td>1.11</td>
<td>$2.90 \times 10^{-01}$</td>
<td>8.65</td>
<td>$3.30 \times 10^{-03}$</td>
<td><strong>4.94</strong></td>
<td>$2.62 \times 10^{-02}$</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>4.32</td>
<td>$3.70 \times 10^{-02}$</td>
<td>**</td>
<td>0.77</td>
</tr>
</tbody>
</table>

$** p < 0.001$, $* * p < 0.01$, $* p < 0.005$

The $\chi^2$ value is the test statistic. Degrees of freedom (df) is the number levels for a given parameter minus 1. The p-value is determined by comparing the test statistic and the df to the $\chi^2$ distribution.

1. In the Storyteller data there were three NP tokens, which all occurred with *have got to*. This data’s full model therefore tests the three-way distinction between generic subjects, specific subjects (NPs and pronouns) with objective obligation, and specific pronominal subjects with subjective obligation. The Storyteller subject/obligation type df is therefore 3. There were no NP tokens with subjective obligation in the Post-Industrial data. The df for Post-Industrial subject/obligation type is therefore 3.

2. In the Storyteller data there were only three female tokens. Sex therefore was not included in the full model for this data.

The log-odds reported for this model have been converted to factor weights so as to be more easily comparable to analyses presented in the sociolinguistic literature. The fixed effects included in each model derive from the results of Wald $\chi^2$ analysis of the full model. The factor weights for non-significant parameters (indicated with square brackets) come from the values reported in the full model (again following the conventions in the

---

sociolinguistic literature). For the Storyteller data the most parsimonious model includes no parameters at all. This is likely due to the very small number of tokens (47); with so few tokens parameters are less likely to be identified as significant over and above chance variation, or the random effect of speaker. Perhaps the best synthesis of the Storyteller data, given that there are less than 50 tokens, is simply that in this data, which is older in both real and apparent time there is a higher rate of non-\textit{have to} variants, and that this high rate is not simply because all of the tokens in the data are in non-favouring environments.

For the Steelworker data the most parsimonious model includes subject/obligation type, year of birth, and sex. For the Post-Industrial data the most parsimonious model includes subject/obligation type, and year of birth. The analysis of the Steelworker and Post-Industrial corpora reveal that both show the same grammatical patterning with respect to subject/obligation type. In both corpora contexts with objective obligation and tokens with generic pronouns (which are almost always objective obligation contexts) favour \textit{have to}, while subjective obligation contexts, no matter what type of subject, disfavour \textit{have to}. In both the Steelworker and Post-Industrial data the difference between NP Objective, Generic Pronoun and Pronoun Objective was not significant. Likewise the contrast between Pronoun Subjective and NP Subjective was not significant. If we re-label generic pronouns as objective, which is logical given that generic pronouns are almost always objective anyway, the relevant parameter is thus clearly Objective vs. Subjective.

Though there were too few tokens to include sex in the analysis of the Storyteller data, the patterning of those few tokens does align with the sex patterning in the Steelworker data. In both corpora women use more \textit{have to} relative to men. This is also the case with the Post-Industrial data, though the difference between men and women did not reach the level of significance. The use of one variant by women more than men is found repeatedly in the sociolinguistic literature when that variant is new, overtly prescribed, or both. Women are generally more advanced than men in the direction of change of the community (e.g., Labov 2001, 266, and elsewhere). For Cape Breton’s deontic modality system this is case, though the difference between men and women among some birth years is not large.

8.7 Discussion

In Inland Canadian English — as exemplified by the data from Toronto presented by Tagliamonte and D’Arcy (2007b) and Southern Ontario presented by Tagliamonte and Denis (2014) — obligation type is not only a significant predictor of the variation, but generic subjects and objective obligation contexts pattern together in favouring \textit{have to}, which is increasing in use in apparent time. In Scottish/Irish English (and Northern British English more generally) — as exemplified by the Roots Archive data presented by Tagliamonte and Smith (2006) and Tagliamonte (2013a) — obligation type is also a significant predictor of the variation; however, generic subjects pattern together with subjective obligation contexts in disfavouring \textit{have to}, which is on the decline in apparent time. In the data presented above for Cape Breton English, obligation type is not only a significant predictor of the variation, but generic subjects pattern together with objective obligation contexts in favouring \textit{have to}, which is on the rise in apparent time. In other words, Cape Breton English aligns with Inland Canadian English and not Scottish/Irish/British English in both frequency of \textit{have to} use among speakers of similar ages, but also the specific grammatical constraints governing the variation between \textit{have to} and other variants.

The \textit{prima facie} conclusion to be drawn from the above results is that use of \textit{have to}, \textit{have got to} and \textit{got to} do not pattern like the use of \textit{have}, \textit{have got} and \textit{got} with respect to subject type. In Section 7.5.3.5 stative possessives showed a very specific pattern with respect to subject type — a pattern that is also only attested in other Canadian data. NP subjects favoured \textit{have} the most, followed by third person plural pronominal subjects,
Table 8.8: Mixed effects logistic regression—contribution of factors to the probability of *have to* for expressing deontic modality across three corpora of Cape Breton English.

<table>
<thead>
<tr>
<th></th>
<th>Storyteller</th>
<th>Steelworker</th>
<th>Post-Industrial</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>aic</strong></td>
<td>45</td>
<td>1912</td>
<td>266</td>
</tr>
<tr>
<td><strong>input</strong></td>
<td>.16</td>
<td>.92</td>
<td>.72</td>
</tr>
<tr>
<td><strong>total N</strong></td>
<td>47</td>
<td>1855</td>
<td>232</td>
</tr>
</tbody>
</table>

**Fixed Effects**

<table>
<thead>
<tr>
<th>Subject/Obligation Type</th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP Objective</td>
<td>KO</td>
<td>0</td>
<td>3</td>
<td>.66</td>
<td>80</td>
<td>135</td>
<td>.62</td>
<td>85</td>
<td>13</td>
</tr>
<tr>
<td>Generic Pronoun</td>
<td>.86</td>
<td>20</td>
<td>15</td>
<td>.54</td>
<td>69</td>
<td>852</td>
<td>.61</td>
<td>80</td>
<td>73</td>
</tr>
<tr>
<td>Pronoun Objective</td>
<td>.50</td>
<td>25</td>
<td>12</td>
<td>.54</td>
<td>72</td>
<td>517</td>
<td>.53</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>Pronoun Subjective</td>
<td>.14</td>
<td>18</td>
<td>17</td>
<td>.40</td>
<td>62</td>
<td>315</td>
<td>.26</td>
<td>43</td>
<td>74</td>
</tr>
<tr>
<td>NP Subjective</td>
<td></td>
<td>0</td>
<td>.36</td>
<td>61</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td><strong>range</strong></td>
<td></td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sex**

<table>
<thead>
<tr>
<th></th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>67</td>
<td>3</td>
<td>.81</td>
<td>96</td>
<td>23</td>
<td>[.57]</td>
<td>67</td>
<td>143</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16</td>
<td>44</td>
<td>.19</td>
<td>69</td>
<td>1832</td>
<td>[.43]</td>
<td>55</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td><strong>range</strong></td>
<td></td>
<td></td>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Year of Birth**

Continuous factor [ ] +0.72 log-odds +0.27 log-odds

**Random Effects**

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Variance</th>
<th>N</th>
<th>Variance</th>
<th>N</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.94</td>
<td>15</td>
<td>1.73</td>
<td>143</td>
<td>2.28</td>
<td>55</td>
</tr>
</tbody>
</table>

Sum contrast coding.

Aic, input, and factor weights for significant factor groups derived from most parsimonious model. Factor weights for non-significant factor groups derived from full model.

Factor weights converted from log-odds. \( p = \frac{e^z}{(1+e^z)} \).

Factor weights with non-significant contrasts indicated with [ ].

Correlation of fixed effects: Steelworker \( r < |0.022| \), Post-Industrial \( r < |0.119| \)

Additional treatment contrast coding used to test orthogonal contrasts for parameters (factor groups) with more than two levels (factors) using most parsimonious model.

Significant contrasts indicated with horizontal lines.
first person pronominal subjects, second person pronominal subjects, and finally, with the least amount of 

have, third person singular pronominal subjects. In Table 8.8 the relevant contrasts are between objective and subject-

ive obligation types, regardless of whether the subject is pronominal or not. This would suggest that whatever phenomena were at work to create the very specific pattern found for subject type in stative possessives may not be at work for deontic modality — a reality that seems implausible given that the forms are the same. If, for example, the contraction of have opens the door for got insertion, which results in pronominal subjects having higher rates of have got/got diachronically, 18 why would this process only occur in stative possessive contexts and not deontic modality contexts — especially when have to is arguably a more grammatically-entrenched expression than have, functioning as a semi-modal? The answer is that it does not. The same process occurs in both the stative possessive and deontic modality systems, but this phenomenon is masked by the much stronger effect of obligation. If we isolate just objective obligation tokens, for which there is robust data, and for which must and need to are inconsequential (see Figure 8.8), we find a similar pattern to stative possessives with noun phrase subjects and third person plural pronominal subjects favouring have to relative to other subject types, followed by first and second person pronominal subjects, which pattern together, and finally third person singular pronominal subjects. A mixed-effects logistic regression analysis, testing the fixed effects of subject type, sex, and year of birth, with speaker as a random effect, on just the objective obligation data found subject type to be a significant predictor of the variation. A secondary model on the same data using treatment contrasts found that the significant contrasts were those represented by dashed lines in Figure 8.12, the same contrasts that can be found for the Steelworker data in Table 7.8. 19

Figure 8.12: Expressions of deontic modality in Cape Breton English. Variants by subject type in objective obligation contexts only. Dashed lines indicate significant contrasts.

18. The specific patterning of Canadian stative possessive systems is believed to result from have and has contracting to ‘ve and ‘s and subsequently inserting got at different rates/times diachronically in Canadian but not other Englishes.

19. This pattern was found when the data from all three corpora were considered together. A conditional inference recursive partitioning testing the combined effects of speaker year of birth, corpus, sex, and subject type found the same pattern among speakers born in the Steelworker and Post-Industrial corpora. Models of these analyses are presented in Appendix A
8.7.1 The lexical status of *have*

Given the similarity between the subject type grammatical conditioning of both *have* and *have to* and their similarity in frequency across the apparent time window (as in Figure 8.13), the patterning of both stative possession and deontic modality in Cape Breton English can, when considered together, provide additional insight into the grammatical system: specifically the status of the verb *have*.

![Figure 8.13: Distribution of stative possessive have and deontic modality have to in Cape Breton English.](image-url)

D’Arcy’s (2015) analysis of stative possession in Victoria’s *the British Colonialist* puts forth the idea that the expansion of *have got*/got forms in Canadian varieties was halted in the 19th century by the indiscriminate expansion of *do*-periphrasis to the incipient *have got* possessive, even though *do*-support is not grammatically necessary in such a construction. She concludes that possessive *have* has reached full lexical status in Canada. She, however, puts forth this idea based on written material that contains very few instances of *have got*/got.

Data from Cape Breton provides strong supporting evidence for D’Arcy’s (2015) proposal. Not only are the case studies of stative possessives and deontic modality expressions the largest from vernacular spoken data, they are also the only studies of vernacular Canadian English with robust use of both the *have got*/have got to and *got*/got to forms. The wholesale reversal of both the stative possessive and deontic modality systems occurred very rapidly in real/apparent time in the Cape Breton data. Among speakers born in the 1880s and in the first decade of the 20th century both *have got* and *have got to* are the majority forms, but within a decade or two *have* and *have to* take over. This rapid reversal is corroborating evidence for D’Arcy’s (2015) proposal that the abandoning of *have got* and *have got to* is a reflex to a different change within the grammar — namely the spread of *do*-periphrasis to lexical *have*.

In keeping with this idea I argue that the reorganization of secondary variants within the stative possessive and deontic modality systems is also evidence of this change. Among speakers born in the 1930s onwards the overall rate of *have got*/got and *have got to*/got to forms gradually decrease, however the relative distribution of these two variants cross over. In other words, *have got*/have got to rapidly decrease, and their position as secondary variants is taken over by the increased use of *got*/got to. The stative possessive and deontic modality

---

20. In the data reported for Toronto by Tagliamonte and D’Arcy (2007b) *got to* also takes over from a dwindling *have got to* in about the 1960s.
systems have rapidly expunged the have got/have got to variants in favour of stigmatized got/got to and neutral have/have to — the two variants for which do-periphrasis is possible. This pattern is exactly what is predicted to occur if the loss of have got/have got to was epiphenomenal to the spread of do-support to a fully-lexical have.

### 8.7.2 Have got to/got to

What about the spike in use of got to among speakers in the 1980s observed in Figure 8.5?

The choice of got to as the main alternative variant to have to by these speakers is, on the one hand, related to the status of have, which is increasingly becoming a full lexical verb in Canadian English as discussed above, even if it is functioning like a semi-modal. However, as with stative possession, there is also evidence from folk linguistic literature (e.g., Gray 2006; 2007) that bare got to is evaluated by speakers as local, even though it was never the most common variant among speakers, and it is this evaluation that is driving up the frequency of its use.

Speakers born in the 1980s represent a particular group of Cape Bretoners. These are people who graduated high school after the collapse of heavy industry in the community. Those born in the 1990s were still high-school students, and younger than 20 when they were interviewed in 2009. This means 1980s-born speakers represent those people who consciously made the choice to stay in Cape Breton and/or those for whom leaving was not possible. This group is in the process of forming their adult identities in a post-industrial Industrial Cape Breton and the signifiers of that identity likely include an increased use of features deemed local (see Section 3.2). Even though the oldest speakers did not have a high rate of got to, instead using mostly have got to, the young Cape Bretoners are using stigmatized got to more often.

### 8.8 Conclusion

Not only does the Cape Breton data show the very specific Canadian-only subject type pattern for stative possessives, it also shows the same very specific Canadian-only subject type pattern for deontic modality. This provides more evidence for a genetic link between Cape Breton English and Inland Canadian English rather than Scottish/Irish or Northern British English. This finding also adds to the literature on the changing deontic modality system. It shows that obligation type has a strong effect over and above the diachronic consequences of grammaticalization that result in the unique Canadian pattern for subject type (see Footnote 18 on p. 119).

Have to shows signs of full entrenchment. This includes taking over for must even in the highly circumscribed contexts where must has held on: metadiscursive pragmatic markers. This suggests that while metadiscursive pragmatic markers resist innovation, change is still possible, perhaps inevitable. It also has elucidated the need for future studies of variation in the deontic modality system to take discourse context into account and to treat metadiscourse and primary discourse separately.

Finally, while need to may be a burgeoning variant in Ontario English (Hoffman and Walker 2010), it is negligible in Cape Breton. This suggests that while the trajectory of deontic modality forms is shared across Loyalist-founded communities, new variants to the system, like need to, can begin in one community and may, in the future, diffuse to others.

Taken together with the findings presented in Chapter 7, the findings from the study of Cape Breton’s deontic modality system point strongly to the transmission of variable patterns from pre-revolutionary American English to both Cape Breton and Ontario, rather than subsequent diffusion of those same patterns from Ontario eastward. Any claims that the similarities between Cape Breton and Ontario’s stative possessive and deontic
modality systems are due to diffusion must account for the preponderance of evidence that there is no simplification of grammatical patterns between these two communities. Changes to both the deontic modality and stative possession systems are occurring via well-documented diffusion in the British Isles, and in the British, Irish, and Scottish communities where these changes are diffusing both a temporal lag and simplification are evident. The argument that rapid diffusion — so rapid to be unobservable — with a lack of grammatical simplification occurred in the 1800s for these two variables is untenable. This is especially true considering these two key features of diffusion are observable for the spread of *be like*, which has occurred during a time of greater inter-dialectal contact (i.e., the last 30 years). In fact, Cape Breton and Southern Ontario contrast with other communities in which *be like* has spread (Tagliamonte, D’Arcy, and Rodríguez-Louro 2016) exactly because they show these hallmarks of diffusion.
Chapter 9

Future Temporal Reference

9.1 Introduction

The final case study is of a morphosyntactic variable that does not involve the forms *have* or *got*, but that can be similarly used in assessing the genetic relationship between Cape Breton English and other varieties. This variable, the expression of future temporal reference, has a long time-depth in the English language and the expansion of *will* and then *be going to* for expressing future time has been tracked across a number of varieties within the paradigms of historical linguistics, variationist sociolinguistics, and corpus linguistics. In the following sections I will comprehensively outline the observations made by these studies, and explore whether the variable future temporal reference system in Cape Breton operates similarly to Inland Canada, Scotland, or another variety.

9.2 The Variable

The English tense system does not include a morphological future tense. Instead, it only distinguishes between past and non-past (Quirk et al. 1985, 176). In order to refer to future time several strategies are possible within the language. These include:

1. Based on the constructions identified by Quirk et al. (1985, 213–218) and Gramley and Patzold (1992, 145), plus *be going to* + present progressive, which Quirk et al. (1985) do not include. This list does not include so called “future in the past,” which is generally expressed with *would* or *was/were going to*, or future modal constructions, which never idiomatically occur with *shall* or *will* in mainstream Canadian English.

(70) a. *shall* + infinitive: *I shall leave tomorrow.*
   b. *shall* + present progressive: *I shall be leaving tomorrow.*
   c. *will* + infinitive: *I will leave tomorrow.*
   d. *will* + present progressive: *I will be leaving tomorrow.*
   e. *be going to* + infinitive: *I am going to leave tomorrow.*
   f. *be going to* + present progressive: *I am going to be leaving tomorrow.*
   g. *be to* + infinitive: *I am to leave tomorrow.*
   h. present progressive + temporal adverb: *I am leaving tomorrow.*
   i. simple present + temporal adverb: *I leave tomorrow.*

In the Cape Breton data *shall*, *will* and *be going to* vastly outnumber other ways of referring to future time.

(71) I literally went to Alberta out, right out of high school. I left in September after graduating in June. Just because one of my closest girlfriends went out there. And I had not made any post-high school plans. I was, I had no direction. And
I said, “Well, I’ll go try it. They’re, they’re going out,” — and, they are out there now — “and I have other friends that are going to go out so I’m going to go.” (Female born 1958, PI)

I told Steve, “I’m not a miner,” and he said, “Shut up. You want to see France, don’t you? First thing we know, the war will be over, and the 64th is never going to get there.” (Male born 1880, ST)

The use of *be going to* is singled out as a local Cape Breton speech feature in the folk dictionary *Da Mudder Tongue*, though it is unclear whether or not it is the use of the morphosyntactic feature or its pronunciation that is considered local given that both future temporal reference *be going to* and present progressive *go* are listed as meanings. Also, the example sentence with a *be going to* future is a so-called “future in the past” construction.

*GOAN TA*— About to perform or do a requested task or heading in a certain direction. “I was goan ta fix that today.” Or, “I was goan ta the store to get some grog.” (Gray 2007, 24)

In other definitions, however, *goan ta* or the more conventional *gonna* form of *be going to* is used.

*GOB*— A term for face taken from Irish descendants.”If she complains once more, I’m goan ta give her a belt in the bob!” (Gray 2007, 24)

*ANNA RISM*— A very risky medical condition when a blood vessel ruptures. “If you don’t keep quiet bye yer gonna give me a anna rism.” (Gray 2007, 3)

In both of the preceding examples *be going to* is used in apodoses — the main (consequent) clauses of a conditional sentence (see Section 9.6.4.4). It is noteworthy that *will* also appears in this exact same syntactic construction in other definitions in *Da Mudder Tongue*, as in (76). Therefore, whether or not the *be going to* future itself or simply its pronunciation is considered a local feature is not fully clear.

*CLIME*— To venture upwards. “If I hafta clime dem stairs once more I’ll have a conniption.” (Gray 2007, 14)

9.3 Historical Perspective

The use of modal *shall* and *will* for indicating future time dates back to Old English (Bradley 1911), while the construction *be going to* — originally with a concrete locative/directional meaning — began to be used for future reference in the late 1400s (Danchev and Kytoö 1994; 2002; Fischer and Wuff 2006, 133–134).

Bradley (1911) points out that in the Middle and Early Modern English periods *shall* and *will* preserved a semantic link to their original use as modals:

The separate colors of *shall* and *will* in this new function were entirely in line with the great tradition of their older service in the modal field. Whatever was regarded as programmed or predestined, was expressed by *shall*. Whatever was to come about through the will or consent of the agent, was expressed with *will*. (Bradley 1911, 18)

In the late 18th and early 19th century the meaning of *shall* expanded to include one of personal compulsion and coercion — thus, when used with second and third person subjects, *shall* acquired an illocutionary force considered too strong and therefore discourteous or rude in cultivated speech. For this reason “I shall” became the sole construction in which *shall* functioned as a neutral future marker. The use of *shall* with first person singular subjects and *will* with other subjects2 in turn becomes a prescriptive rule (see Boyd and Boyd 1980, for a summary of *will* and *shall* prescription).

2. This pattern of first person singular *shall* but otherwise *will* as a neutral future was required to be reversed for interrogative or conditional phrases, according to prescriptive grammars of the time (e.g., Murray 1817, 98–99)
Most prescriptive texts point out that *shall* was seldom used correctly — if used at all — in rural Britain, Scotland, Ireland, or North America, other than by those educated in British norms.

I never knew an Englishman who misplaced “shall” and “will.” I hardly ever have known an Irishman or a Scotchman who did not misplace them sometimes. And it is strange to observe how incurable the propensity is, (Alford 1864, 154).

In New England it is noteworthy that even the boys and girls use *shall* and *will* correctly; and in New York, New Jersey, and Ohio, in Virginia, Maryland, and South Carolina, fairly educated people of English stock do the same; while by Scotchmen and Irishmen, even when they are professionally men of letters, and by the great mass of people of the Western and South-western States, the words are used without discrimination, or, if discrimination is attempted, *will* is given in the place of *shall*, and *vice versa*, (White 1870, 264).

Knowing the abstruse prescriptions for the proper use of *shall* and *will* — L. Bloomfield (1935/1996, 7) calls them “fanciful”— “becomes an index of belonging to the proper class or region” (Boyd and Boyd 1980, 46). Fowler (2009, 526), in laying out the rules for the use of *shall* and *will*, even provides ample examples from “newspapers of the better sort” in which the “principle of idiom is outraged” so that “the Scotch, Irish, provincial, or extra-British writer will thus have before him a conspectus of the pitfalls that are most to be feared.” In other words, he provides examples of the misuse of *shall* and *will* to all those outside of urban Britain by showcasing mistakes in *shall* and *will* use made in reputable British newspapers. The correct use of “I shall”, or the use of *shall* at all, therefore clearly was not universal.

The corollary of *shall*’s strong prescriptive and regional (e.g., south east British English) status in the 18th to early 20th century — the time period in which Cape Breton’s permanent English-speaking settlements were established — is that in vernacular varieties *will* was dominant. *Be going to*, whose meaning broadened from simply movement towards a goal in the 17th century (Pertejo 1999, 136) was also gaining ground as a marker of future temporal reference in the 19th century (Nesselhauf 2010, 170, Figure 2).

As with the putative meaning distinction between *shall* and *will*, prescriptivists have argued that *be going to* and *will* also denote different types of futures. Celce-Murcia and Olshtain (2000) summarize these arguments as follows:

A frequent and persistent question from ESL/EFL teachers and students is “What is the difference between *be going to* and *will*?” Previous accounts, based on the work of Binnik (1971) and McCarthy and Carter (1994), among others, have offered something like this: *be going to* is more informal, immediate, and interactive than *will*, which is more neutral or formal; *be going to* primarily refers to the immediate future or plans in future time, but *will* has several other meanings that are modal in nature, such as promising, predicting with certainty and so forth. (Celce-Murcia and Olshtain 2000, 226)

Descriptive linguistics have, however, disagreed with these claims.

In most cases there is no demonstrable difference between *will/shall* and *be going to* though many scholars have looked without success for one. (Palmer 1979, 163, see also Palmer 1987, 148; Hall and Hall 1970; Quirk et al. 1985, 228; Danchev and Kytoö 1994, 384, and compare with Gramley and Pätzold 1992, 145).

---

3. Nesselhauf’s (2010) analysis of the British part of the ARCHER corpus from 1650 to 1999 shows a steady increase in the use of *will* and *be going to* at the expense of *shall* — though she also considers the rise and fall of other future temporal reference constructions, like those presented in (70).
Mair (1997, 1538) argues that the only difference between be going to and will or shall is stylistic (though see the discussion in Poplack and Tagliamonte 2000, 321).

The going-to future has not become more frequent because it has added any new functions or uses; rather it has become more frequent because it tends to be chosen more often in contexts in which it has long been established and competing with will/shall and other expressions of futurity, for example as the stylistically informal alternative. (Mair 1997, 1541)

The association between be going to and informal styles is echoed in the results of corpus-based studies (Mair 1997; Berglund 2005; Szmrecsanyi 2003, etc.).

Given the usage history and the history of prescriptivism of the future temporal reference system, several questions arise with respect to its status in vernacular Cape Breton English. To what extent does shall play a role in the system? Given that Cape Breton English represents the amalgamation of Loyalist New England English (shall-using, see above) and Scottish and Irish English (both shall-less), and is considered to be highly vernacular, will there be relatively more or less shall than, say, Ontario English? Further, as a vernacular variety (with possibly more historic shall use) to what extent has innovative, but also “informal”, be going to penetrated Cape Breton English?

### 9.3.1 The McKinnon Texts

Unlike the previous two chapters, where the McKinnon Texts offered few tokens and little variation, for future temporal reference there were many tokens in the McKinnon Texts and there is robust variation between shall and will. Across the McKinnon Texts will is the most common modal used for referencing future time (83 percent $N = 359$) in quoted dialogue.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>shall</th>
<th>will</th>
</tr>
</thead>
<tbody>
<tr>
<td>McKinnon Texts</td>
<td>74</td>
<td>17</td>
</tr>
</tbody>
</table>

Of note, shall was more frequent when the subject was first person: 24 percent, ($N = 213$) compared to 10 percent ($N = 230$) for other subjects. This aligns with the prescriptive tradition of shall being used with first person singular subjects — though here all first person subjects, both singular (21 percent, $N = 182$) and plural (39 percent $N = 31$), show higher rates of shall. Shall is also used more frequently with stative verbs like be and have (22 percent, $N = 147$), compared to dynamic verbs like dance, arrive, and make (14 percent, $N = 296$).4

If the McKinnon texts are at all reflective of the community grammar in Cape Breton in the mid-19th century, then we might expect will to continue to be favoured by non-first-person subjects and dynamic verbs relative to shall. While there are no be going to futures in the McKinnon texts, they do appear — along with will and shall — in the works of Thomas Chandler Haliburton and Thomas McCulloch, indicating the be going to future was likely used in Nova Scotia in the 19th century.

---

4. These differences were confirmed to be statistically significant using a series of multivariate linear regression models calculated in R. As there were no random effects for speaker, these models only tested fixed effects. The most parsimonious model — that is, the model that was the most explanatory of the variation in the data — included only subject number (first versus other) and event type (state vs other). Predictors of the variation found to be non-explanatory included polarity, proximity, sentence/clause type, subject animacy, and presence of a temporal adverb — all factors known to constrain variation in the future temporal reference system (though usually variation between will and be going to) in spoken data. See 9.6.4.7.
When writing as the Loyalist Mephisbosheth Stepsure, the Scottish McCulloch employs shall, will, and be going to, but when writing in Scots as Alexander Scantocreesch he only employs ’ll and will. As will be discussed in the next section, this pattern is somewhat at odds with the contemporary distribution of these variants in peripheral varieties of Scottish English.

9.4 Synchronic Perspective

The future temporal reference system has been studied among several groups of Canadian English speakers. Poplack and Tagliamonte (2000) examine the use of future temporal reference variants across four Canadian communities, three of which are in Nova Scotia.\(^5\) They find that the overall rate of use of be going to, and its overall likelihood, was greater in the two African–Nova Scotian Loyalist communities (North Preston, 52 percent, and Guysborough Enclave, 48 percent) than in Guysborough Village (34 percent), also in Nova Scotia, but with a mixed British/Loyalist-origin white population, or Ottawa (46 percent), which the authors use as a proxy for mainstream Canadian English. While future temporal reference variation had been studied and discussed by historical/corpus linguists (e.g., F. T. Visser 1963–1973; Mair 1997; Berglund 1997; Peregrine 1999), Poplack and Tagliamonte (2000) were the first to apply quantitative variationist sociolinguistic methods to examining the

\(^5\) Poplack and Tagliamonte’s (2000) study also includes the English-speaking community of Samaná, Dominican Republic, which was settled in the 1820s by African Americans. The overall rate of be going to in this community is 57 percent. The overall likelihood of be going to is .59, and the two significant constraints on variation between forms is point of reference (future in the past contexts favour be going to) and verb type (motion verbs disfavour be going to).
variation between forms in spoken data. Their analysis included future forms where the point of reference was either the present, as in (79) a., or the past, as in (79) b.

(79) a. An engineer might be able to put in a new piece of equipment and say that it is going to work. (Male born 1927, SW)
b. Usually a guy got it a particular time because he knew his name was going to be called, and he knew the foreman. (Male born 1928, SW)

One of Poplack and Tagliamonte's (2000) major findings was that in each of the communities they examined, future in the past contexts significantly favour be going to. Subsequent variationist analyses of future temporal reference have not included future in the past contexts, mainly because be going to is nearly categorical, but also because the full suite of future options are not grammatically possible in this context (e.g., shall, present, present progressive). Longitudinal patterns for future form use in this context are therefore expected to be different because the suite of available variants is different. The overall rate of be going to use in Poplack and Tagliamonte (2000) is therefore calculated using an extra, highly-favouring context for be going to; comparisons of subsequent analyses to Poplack and Tagliamonte (2000) must take this into account.

The frequency of be going to use among African–Nova Scotians is particularly germane to the study of the roots of Cape Breton’s future temporal reference system. African–Nova Scotians are unquestionably of Loyalist origin (unlike, for example, the white mixed-British/Loyalist-origin Guysborough Village). While this community has shown strong similarities with other AAVE varieties (Poplack and Tagliamonte 1991), and regardless of from where the original roots of AAVE derive, African–Nova Scotians still represent part of the spectrum of Loyalist English that was transported to Nova Scotia. The communities’ conservatism, due to linguistic isolation and resistance to contact-induced change, makes them particularly useful for assessing earlier vernacular forms of Nova Scotia English, especially in light of the fact that our understanding of Nova Scotia’s input Englishes comes from written records and dialect writing.

Keeping in mind that the variable context in Poplack and Tagliamonte (2000) is different from subsequent studies, and from the study presented in Section 9.6.4.7, it is notable that there were some differences between the grammatical patterning of the African–Nova Scotian speakers and the speakers from Guysborough Village and Ottawa. In North Preston and Guysborough Enclave subordinate clauses favoured be going to, while in Guysborough Village and Ottawa there was no difference between clause types. In the Nova Scotian communities, but not Ottawa, verbs of motion disfavoured be going to, while in Ottawa there was no difference between motion and non-motion verbs. Finally, all communities showed higher rates of be going to with non-first person subjects, but this effect was only significant in Guysborough Enclave and Ottawa. Both Guysborough Village and Ottawa showed a stronger distinction between immediate futures (which favoured be going to) and distal or non-specified futures, which favoured other forms. The African–Nova Scotian communities did not show this clear distinction. The authors concluded that all of the Nova Scotian communities represented relic forms of North American English and older configurations of the future temporal reference system. Speakers from Guysborough Village did, however, pattern only in some ways like the more-relic African–Nova Scotian communities (with verb type effects), while in other ways more like the mainstream Ottawa speakers (with proximity effects). The Guysborough Village speakers therefore likely represent a transitional stage in the development of future be going to.

The corpus of 82 Nova Scotia speakers used by Poplack and Tagliamonte (2000) was collected in 1990-1991 (Poplack and Tagliamonte 1991). All of the speakers were over the age of 53, which means their birth years were

6. Though by 1999, Labov and his collaborators (Labov et al. 1968, Labov 1972, and others) had examined the use of gonna as a form of phonetic reduction in African American Vernacular English within the sociolinguistic paradigm.

7. For example, the rate of be going to for Ottawa, with future in the past contexts removed, is 27 percent. This value was reported in a footnote of Torres Cacoullos and Walker (2009, Footnote 11), based on a personal communication between the authors of both studies.
sometime between 1897 and 1937. Their birth years therefore align with the birth years in the *Storyteller* corpus and the oldest speakers in the *Steelworker* and *Post-Industrial* corpora, while the time of collection aligns with that of the *Steelworker* corpus (1989/1990). The 17 speakers from Ottawa in Poplack and Tagliamonte (2000) were age-matched to those from Nova Scotia, so they were also born sometime in the first four decades of the twentieth century. The speakers are a subset of anglophones from the *Ottawa-Hull Spoken Language Archive*, which was collected between 1983 and 1998.

Torres Cacoullos and Walker (2009) explore the role of collocation effects in the temporal reference system in their study of 74 speakers from Québec City and Montréal, recorded in 2002–2003 (see Poplack, Walker, and Malcolmson 2006). Their analysis excluded all *shall*, future in the past, and certain temporal clauses headed by *when*, *until*, and *before*. The authors found an overall rate of 51 percent *be going to* in their data. This is nearly double the rate found for Ottawa by Poplack and Tagliamonte (2000) (based on the rate found by the authors minus future in the past contexts), but Torres Cacoullos and Walker’s (2009) data includes younger speakers and was collected between five and twenty years later. Torres Cacoullos and Walker (2009) do not operationalize age in their study, so it is unclear whether the increased rate they find is a regional difference or representative of Québec City and Montréal being further along a trajectory of change — something that would have been elucidated by showing apparent time change in their dataset.

The authors find a strong favouring of *will* with first person subjects, especially with the expression *I’ll tell you...* and when statements were negated with *never*. The authors find other highly circumscribed contexts where one variant or the other was highly favoured, which they concluded was the result of each variant retaining traces of its initial entry point of grammaticalization as a future marker. Or as Tagliamonte (2013b, 148) summarizes Torres Cacoullos and Walker’s (2009) findings, “extant variants remain entrenched in certain contexts rather than undergoing extension.” From a general perspective, however, Torres Cacoullos and Walker’s (2009) findings do adhere to the patterns found in Nova Scotia and Ottawa of *be going to* being favoured in subordinate clauses and with non-first person subjects. They also found that the variant was favoured in questions.

Tagliamonte and D’Arcy (2009) examine the rise of *be going to* in the *Toronto English Archive*. They conclude that the rise of *be going to* may be progressing very slowly or even slowing down. The also hypothesize that the division of labour between *be going to* and *will* could be solidifying. The authors find a borderline significant apparent time increase in *be going to* use among their informants, rising from 39 percent among speakers older than 60 to about 48 percent among young speakers — though teenage female speakers used *be going to* about 58 percent of the time. While the overall growth from oldest to youngest was significant, there was no systematic change by age cohort (i.e., the difference between young speakers and older speakers was large enough to reach the threshold for statistical significance, but the difference between young speakers and middle-age speakers, or middle-age speakers and old speakers was not large enough to be statistically significant). Further, there was no significant sex effect for any age group, even though among the same speakers, for other variables — including variables with long time depths like changes to deontic modality and stative possession — there is sex differentiation. This is in part why the authors claim that *will* and *be going to* may be stabilizing into functional partitioning. And as Tagliamonte (2013b) points out, if a change has levelled out the size of the gradient effect would be small.

Tagliamonte and D’Arcy (2009) include all examples where the reference time is after the speech time, but exclude all non-temporal contexts (including hypotheticals and counterfactuals, etc.,) as well as future in the past. Like Torres Cacoullos and Walker (2009), the authors find that *be going to* is highly correlated with questions,

---

8. These were invariably simple present futures.

9. Torres Cacoullos and Walker (2009, 345) argue that *be going to* and *will* are actually at advanced stages of grammaticalization because they appear in both of these clause types.
while *will* is highly correlated with *never* negation, and more likely than *be going to* with first person subjects (detailed in Tagliamonte 2013b). They argue that while Torres Cacoulllos and Walker (2009) might claim that variants remain entrenched in certain highly circumscribed contexts, other more substantial contexts do not appear to show the same extreme imbalance towards one variant or the other.

Denis and Tagliamonte (2017) re-examine the future temporal reference data presented by Tagliamonte and D’Arcy (2009), with an eye to testing the latter authors’ hypothesis that *be going to* and *will* are stabilizing into functional partitioning. Denis and Tagliamonte (2017) find that statistically significant, albeit very slow, change does occur over apparent time in the Toronto data, and while *will* and *be going to* are specialized in some linguistic contexts (apodosis for *will*, questions for *be going to*, etc.), there is a more general pattern of a constant rate of increase in the use of *be going to* across other contexts — indicating grammatical competition.

In British English the most comprehensive analysis of future temporal reference variation in a spoken dialectal corpus is Tagliamonte, Durham, and Smith (2014). The authors combine speech data from studies across the British Isles collected between 1997 and 2009, including the peripheral Northern Irish and Scottish communities discussed in earlier chapters. To these communities, the authors add Lerwick in the Shetland Islands off northern Scotland.

They conclude, based on their findings, that *shall* is rare or non-existent in Scotland and Northern Ireland (a pattern reported elsewhere by Crystal 1986, 42; Filppula 2012; Corbett and Stuart-Smith 2012, 89; Bann 2013, 57 etc., though future *shall* did exist in Scots, Purves 2002, 51). The authors found that the internal constraints for *be going to* use were different depending on how vigorous *be going to* was used among speakers. Younger speakers and speakers from less peripheral, southern communities used *be going to* the most vigorously. The oldest speakers from the most peripheral communities, e.g., Scotland and Northern Ireland, used *be going to* to the least.

Across the full dataset *be going to* is favoured in subordinate clauses, but other constraints were wildly different. For example, in the most conservative varieties *be going to* was most likely in interrogative, negative polarity and first person intentional contexts. They argue this is because these contexts must be part of the triggering environments for *be going to*’s grammaticalization (cf. Hopper and Traugott 1993, 88). In varieties where *be going to* is used more vigorously new patterns emerge including a resistance to *be going to* with first person subjects (perhaps because of prescriptive norms or the collocation *I’ll*), and the extension of *be going to* to non-animate subjects and far-future events.

It is worth noting that the conservative Scottish and Irish varieties in Figure 9.1 are also extremely conservative across a number of ongoing change in progress (e.g., Tagliamonte 2013a; Tagliamonte, Smith, and Lawrence 2005; Tagliamonte and Smith 2006, etc.). Tagliamonte, Durham, and Smith (2014) categorize *be going to* as “incipient” among the oldest speakers (based on Nevalainen and Raumolin-Brunberg 1996, 213–55) because it is used less than 15 percent of the time. They categorize *be going to* as “new and vigorous” among the younger speakers, and among the speakers from more southern communities in their study, because it occurs between 15 and 36 percent of the time.

---

10. Future temporal reference data from these communities is also reported in Tagliamonte (2013a).

11. This finding is at odds with the use of *shall* by Scottish-born and educated Thomas McCulloch outlined above. (McCulloch graduated in logic from the University of Glasgow in 1792 and spoke several languages.) McCulloch employs *shall* frequently, and not exclusively in the prescriptively appropriate *I shall* context. This may be because he is mimicking Loyalist speech, but this seems unlikely given that he elsewhere employs informal vernacular forms for these characters and *shall* was considered formal. It may be because it is prescriptively appropriate; however, he employs *shall* outside of the prescriptive contexts, e.g., with third person subjects. This suggests that *shall* may be part of his own vernacular usage — something unexpected of a Scot. When McCulloch writes in Scots he does not employ *shall*, even though the form existed in Scots. If McCulloch’s use of *shall* is representative of Scottish English speakers in the 19th century, it is possible the form may have been at least a stylistic option that either died out before it could be attested or was infelicitous in the contexts in which the above studies drew their data.
Shall is also rare in other transported varieties, including Newfoundland English (Clarke 2010, 89); Australian English (Collins 2014), New Zealand English (Hay, Maclagan, and Gordon 2008, 52; Gordon 2012, 329), and Trinidadian, Jamaican, Bahamian, Fijian, Indian and Singapore English (Deuber et al. 2012), where will and be going to are more common. In New Zealand, be going to has grown from relatively rare 100 years ago, to used about 50 percent of the time (Hay, Maclagan, and Gordon 2008, 52)— consistent with rates in Canadian varieties. Collins (2014), who examines lexical frequency, finds will is more than three times more frequent than be going to in Australian fiction written since the beginning of the 19th century; however, be going to increases from about < 0.18 instances per thousand words (ptw) in the 1800s to about 0.57 ptw in the 1900s, while will decreases from about 3.6 in the 1800s ptw to about 2 ptw in the 1900s, suggesting a change in progress. Deuber et al.’s (2012, 89) analysis of future expressions in samples of Trinidadian, Jamaican, Bahamian, Fijian, Indian and Singaporean English finds all six varieties use be going to/ gonna/ gonna/go/ a go/ gon’ between 26 percent (India) and 56 percent (the Bahamas) of the time — though most varieties have a rate of 35−41 percent.

Of the number of corpus-based enquiries on the future temporal reference system, a general pattern of more be going to in American English data samples compared to British English data samples has been found (Wekker 1976; Berglund 1997; Szmrecsanyi 2003; Collins 2014).

### 9.5 Methods

As with the preceding chapters the aim of this case study is to make the results as maximally comparable to past studies as possible. Thus the methods for extracting and coding examples of future temporal reference from the Storyteller, Steelworker and Post-Industrial corpora was modelled on past studies, including Tagliamonte and D’Arcy (2009), Torres Cacoullos and Walker (2009), Tagliamonte, Durham, and Smith (2014), and Denis and

---

12. The authors categorize gon’ as creole future marker (Deuber et al. 2012, 88, 97) and thus differently from be going to. This form only occurs in the Bahamian English sample which is has a rate of 35 percent be going to/ gonna and a rate of 31 percent gon’, which combined equal 56 percent. Go occurred once in the Trinidadian data; a go occurred once in the Jamaican data.
Tagliamonte (2017). Each instance of present tense *shall, will,* and *be going to* with future temporal reference were extracted from the data. So-called “future in the past” contexts were not extracted because past studies have deemed this context to fall outside of the envelope of variation — as noted above, *shall* and *will* are infelicitous in this context — and excluded it. Other markers of future time were extracted (e.g., simple present with temporal adverbs) but were not included in the present analysis because they were extremely infrequent relative to *will* and *be going to* and because they are distributionally and likely semantically distinct from *will* and *be going to* (F. T. Visser 1963–1973; Huddleston and Pullum 2002, 131–133; Gramley and Pätzold 1992, 145; Copley 2009, 15–16; Denis and Tagliamonte 2017, 3). All examples of ‘ll were considered to be *will* rather than *shall* following Aarts, Close, and Wallis (2013).

### 9.6 Results

#### 9.6.1 Trajectories of Change

Over the last 130 years there has been a gradual rise in the use of *be going to* as a marker of future temporal reference, going from about 25 percent among the speakers born in at the turn of the 20th century, to about 52 percent among speakers born at the turn of the 21st.

![Graph showing trajectories of change in expression of future temporal reference across corpora of Cape Breton English](image)

The overall rates of *be going to* use in the three Cape Breton corpora meet the threshold for classification as “new and vigorous” according to Nevalainen and Raumolin-Brunberg (1996), unlike the rates reported for older speakers in the Scottish and Northern Irish corpora by Tagliamonte, Durham, and Smith (2014). In fact, at 44 and 42 percent, the rate of *be going to* in the Steelworker and Post-Industrial corpora qualify as “mid-range” changes (i.e., 36–65 percent). The rates of *be going to* in the Cape Breton data more closely match the results reported for other peripheral North American varieties, including the African–Nova Scotian communities and Ottawa reported by Poplack and Tagliamonte (2000), Toronto reported by Tagliamonte and D’Arcy (2009), and the Caribbean varieties reported by Deuber et al. (2012, 89).
Table 9.2: Distribution of expressions of future temporal reference in Cape Breton English.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>shall</th>
<th>will</th>
<th>be going to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Storyteller</td>
<td>1</td>
<td>0.2</td>
<td>395</td>
</tr>
<tr>
<td>Steelworker</td>
<td>3</td>
<td>0.1</td>
<td>1,998</td>
</tr>
<tr>
<td>Post-Industrial</td>
<td>1</td>
<td>0.2</td>
<td>340</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>0.1</td>
<td>2,733</td>
</tr>
</tbody>
</table>

Figure 9.3: Distribution of *be going to* across varieties of English. Adapted and expanded from Denis and Tagliamonte (2017, Figure 5). Purple bars represent values from speakers older than 60 within a community. Orange bars represent values for a whole community. The two green bars represent the two older Cape Breton cohorts: the *Storyteller* corpus (ST), and the speakers from the *Steelworker/Post-Industrial* corpora (SW/PI) born before 1949. Data from communities marked with an asterisk and with light orange bars include the highly-favouring "future in the past" environment, so their values must be considered *cum grano salis*. Values are derived from Poplack and Tagliamonte (2000, Samaná, Ottawa, Guysborough Enclave, North Preston, Guysborough Village), Deuber et al. (2012, Bahamas, Trinidad, Jamaica, Singapore, India), Torres Cacoullos and Walker (2009, Québec City and Montréal), Hay, Maclagan, and Gordon (2008, New Zealand), Szmrecsanyi (2003, *Corpus of Standard North American English, British National Corpus* (DS)), Tagliamonte and D’Arcy (2009, Toronto), and Tagliamonte, Durham, and Smith (2014, Wheatley Hill, Tiverton, Maryport, York, Lerwick, Henfield, Cumnock, Buckie, Northern Ireland). Critical values for changes in progress based on Nevalainen and Raumolin-Brunberg (1996).
Figure 9.3 shows how Cape Breton be going to use compares to rates reported for other varieties. The Cape Breton data (the green bars) was divided into two groups: the Storyteller corpus and a mixed Steelworker/Post-Industrial cohort that includes only speakers born before 1949. Grouping speakers this way is based on the analysis presented in Section 9.6.4.7.

The older, more rural speakers from the Storyteller corpus are the transported variety of English with the lowest rate of be going to. They employ less be going to than was recorded in the British National Corpus. Yet this oldest Cape Breton cohort employs more be going to than all of the communities examined by Tagliamonte (2013a) and Tagliamonte, Durham, and Smith (2014) — including the city of York. Importantly, also, the storyteller data is older than any of the data examined by Tagliamonte (2013a) and Tagliamonte, Durham, and Smith (2014), both in apparent time (the Storyteller corpus includes speakers born in earlier years) and in real time (the Storyteller corpus was collected earlier).

Based just on the rates of use of be going to, it would be safe to conclude that Cape Breton English is a transported variety with vigorous competition between forms. The oldest speakers in the community do use be going to the least frequently, but they employ it more than conservative British/Irish/Scottish varieties. Given that Cape Breton is a peripheral community, and therefore likely to have somewhat decelerated changes, and given that be going to was part of the English language prior to Cape Breton’s English-speaking settlement (meaning all English varieties would have some level of the variant), the above results suggest that the English speaking input to Cape Breton must have been greater than Scottish English could have provided. Also, given that Toronto speakers of a similar age use less be going to — and that Toronto shows a significant increase in be going to use over apparent time (suggesting a trajectory of change in which be going to was used less frequently in the past) — it is unlikely that the rate of be going to in Cape Breton results from diffusion from Inland Canadian English.

That being said, the question remains as to whether the grammatical constraints governing the variation between will and be going to align with those identified for North American or for British/Irish/Scottish varieties. In order to test the combined effects of multiple grammatical factors and determine the exact grammatical patterning of be going to in Cape Breton English, a multivariate regression analysis was employed.

### 9.6.2 Shall

While shall occurred in the McKinnon Texts with some frequency, there were only five instances of future temporal reference shall across the spoken Cape Breton English corpora. They are listed in (80). Given the rarity of shall in the data, it was excluded from further analysis. Four of the five instances of shall were employed by speakers born in the first two decades of the 20th century and all examples were used outside of the prescribed “first person singular subject for positive polarity, declarative sentences” context. The token from the youngest speaker, born in 1958, is used as the perlocutionary marker shall we say — a context, as described below and in Chapter 8, where moribund features tend to be preserved longer diachronically. The token from the speaker born in 1924 is interrogative, but pragmatically is intended to make an offer or suggestion — in this sense, like a perlocutionary marker, it is being employed to organize discourse. The two tokens from the speaker born in 1927 are actually that speaker reciting something written on his job description at the steel plant. While shall may have disappeared from contemporary vernacular speech, Williams (2013) points out that it is still used in legalistic and formulaic expressions. Given these facts, we are left with perhaps a single token of shall in primary, non-formulaic discourse across the 150 years of apparent time encompassed by the Storyteller, Steelworker and Post-Industrial corpora.
9.6.3 Will and 'll

Across the dataset 84 percent of will tokens were 'll. This is in large part because 89 percent of will tokens occurred with a pronominal subject, and pronominal subjects highly favour contraction (L. MacKenzie 2013, etc.). Whether or not a subject was pronominal or an NP did not result in much difference in choice between be going to or a form of will; NP’s, for example, occurred with will 53 percent of the time, while pronominal subjects occurred with will 56 percent of the time. However, this difference in subject type does make a large difference for whether a will token will be contracted or not. Among just the will tokens, 'll is used with NP subjects 22 percent of the time (Total N = 245), but is used with pronominal subjects 91 percent of the time (Total N = 2112). This pattern is in line with research on will contraction (L. MacKenzie 2013, Figures 1, 2).

9.6.4 Linguistic Conditioning

In the following section the numerous linguistic constraints found to (or believed to) constrain the variation between variants in the future temporal reference system will be examined systematically. Several competing findings/theories explaining the variation between forms will be evaluated and a specific coding strategy for testing these findings/theories will be outlined.

9.6.4.1 Sentence Type

In several studies of future temporal reference the greatest predictor of whether will or be going to occurs is sentence type. Negative and interrogative sentences are particularly favourable to grammaticalizing be going to (Berglund 1997; Szmrecsanyi 2003; Torres Cacoullos and Walker 2009; Tagliamonte, Durham, and Smith 2014). There is, however, some evidence that British English and American English differ here. While both varieties show early and consistent favourability of be going to in interrogative contexts, British English appears to favour I won’t or I’ll not rather than I’m not going to for negative sentences (Szmrecsanyi 2003). Among the oldest and most peripheral speakers (Scottish/Irish) in Tagliamonte, Durham, and Smith’s (2014) data, for whom be going to is only incipient, both interrogative and negative sentences favour be going to and this constraint has the highest magnitude of effect. Over apparent time, and in less peripheral communities, however, this pattern changes. Interrogative sentences continue to favour be going to, but whether or not a sentence is negative becomes irrelevant in predicting the variation between be going to and will — effectively matching the pattern observed in the British National Corpus (Szmrecsanyi 2003). This leads the authors to conclude that their data from older Scottish/Irish speakers captures a earlier stage of grammaticalization of be going to in a British English context.

13. It is perhaps ironic, then, that Orkin (1970, 153) labels I won’t an Americanism that Canadians are adopting in lieu of the British I shan’t. Orkin (1970), presumably synthesizing Avis (1955) and Hamilton (1955), reports that four-fifths of questioned Ontarians use I won’t over I shan’t.
Like younger/less-peripheral British English speakers, the speakers from Québec City and Montréal analyzed by Torres Cacoullos and Walker (2009) favour be going to with questions, especially Wh-questions (Yes/No questions 71 percent, Wh-questions 92 percent), but the difference between affirmative declarative (48 percent) and negative declarative (52 percent) sentences appears to be inconsequential.

Tagliamonte, Durham, and Smith (2014, 90) report that a reanalysis of the data presented in Poplack and Tagliamonte (2000) (which includes North Preston, Guysborough Enclave and Guysborough Village, Nova Scotia, and Ottawa) revealed that interrogatives "consistently and strongly favour be going to across all varieties studied." The authors do not indicate whether they found negatives to also favour be going to.

In order to compare the Cape Breton data to previous studies, each token was coded for sentence type, as in (81). In the final regression analysis Yes/No questions and Wh-questions were combined as they did not pattern independently.

(81)

Yes/No Question What are they going to hang us on? Are they going to be trying to fire us and all that? (Male born 1941, PI)
Wh-Question I always ask the question: “When is it going to be that you can’t give your child more than you had?” (Female born 1966, PI)
Affirmative Statement No, she’ll be home December fourth. She was over there for an international conference of some kind. (Female born 1920, PI)
Negative Statement Your car is not going to work very well, it’s not going to be efficient. If it’s not efficient, you can’t afford to run it. (Male born 1945, SW)
never Coal going out, Steel going out— gee whiz — you’ll never see that again. I hate to say it. (Male born 1915, SW)

Distribution of be going to by sentence type in Cape Breton English

Figure 9.4: Distribution of be going to by sentence type in Cape Breton English.

Torres Cacoullos and Walker (2009) identify one confounding issue with respect to negation and future temporal reference, which may or may not have impacted past findings. The temporal adverb never is highly correlated with the use of will. This phenomenon is equally active in the Cape Breton data (see also Section 9.6.4.5). Differences in be going to use in negative contexts across datasets may be the result of this interaction.

14. Sentence type is a significant predictor in Torres Cacoullos and Walker (2009), but the authors do not test the significance of internal contrasts for factor groups with more than two factors.
as some datasets may have more or less never-negation relative to other negation strategies (not, no, etc.) and/or researchers may have made different decisions about how to code it. In the Cape Breton data, non-never negative statements occur with will about half the time (51 percent, Total $N = 529$), but never negated statements occur with will 91 percent of the time.

As in other communities, interrogative contexts show a higher rate of be going to use (88 percent) in Cape Breton English. Also, as in other communities where be going to is fully-entrenched, and no longer incipient, the frequency of be going to in both affirmative and negative statements is roughly the same in Cape Breton English — though, the difference between 49 percent for non-never negative statements and 41 percent for affirmative statements is somewhat greater than the same difference reported for Québec City and Montréal.

### 9.6.4.2 Proximity

The effect of proximity of the reference event to speech time has been variably reported as operational among varieties of English. Generally, go is thought to be favoured when the event time is in the immediate future (e.g., Beattie 1788, 219–220; Quirk et al. 1985, 214). An association between be going to and the immediate future also parallels the grammaticalization of go-like verbs in Romance languages (e.g. Fleischman 1982; Vet 1993; Blas Arroyo 2008). This correlation between be going to and proximity was significant in Poplack and Tagliamonte’s (1991) analysis of British-origin speakers from Guysborough Village, Nova Scotia, though no correlation was found for the speakers in the study who were of African-Nova Scotian-origin (from Guysborough Enclave and North Preston, Nova Scotia), or similarly British-origin speakers from Ottawa. Roy (2007) finds a significant correlation between proximity and be going to in his analysis of 18th-century English writing. Torres Cacoullos and Walker (2009) also find a borderline significant correlation in their Québec City and Montréal English data, with be going to favoured in proximate contexts (even though it was used in proximate contexts 45 percent of the time and in distal contexts 46 percent of the time). The authors attribute not finding a stronger proximity effect to an interaction between proximity, subject type, and adverbial type in their data, which they further suggest may be due to collocation effects. Some common expressions, like *I’ll tell you…*, bolster the frequency of will for first person singular pronominal subjects and for proximate contexts.

Tagliamonte, Durham, and Smith (2014) find proximity effects in their analysis of future temporal reference in their combined corpus of British, Irish and Scottish varieties; however, they authors re-frame the constraint as a resistance to be going to in specifically distal contexts rather than favourability of be like in specifically proximate contexts. They make this claim because tokens for which proximity is not specified also favour be going to.

The Cape Breton data was divided into four groups based on proximity. First, all tokens where the reference event was specified or was contextually understood to occur within 24 hours of the speech event were coded as proximate. Second, tokens where the reference event was specified or was contextually understood to occur 24 hours after the speech event were coded as distal. Third, tokens where the exact or even approximate time of reference was unspecifed — other than just being sometime in the future — were coded as unspecified. Finally, any tokens where the future temporal reference expression was part of a perlocutionary marker (see Chapter 8, Section 8.6.1) were labeled as metadiscourse. This final category captures all of the *I’ll tell you…* tokens, as well as other contexts where the speaker is overtly commenting on the current discourse.

---

15. Szmrecsanyi (2003) only includes NOT-negation (Tottie 1991; Quirk et al. 1985, 782; Cheshire 1999), which includes suppletive won’t, but therefore does not include never negation — or any other NO negation strategy for that matter. Tagliamonte, Durham, and Smith (2014) do not detail whether any negation types were excluded. Poplack and Tagliamonte (2000) do not test sentence type, and Tagliamonte, Durham, and Smith’s (2014) re-analysis of this data does not comment on negation.

16. As Denis and Tagliamonte (2017, 14) point out this contrast is common in languages that distinguish temporal proximity morphologically (either in future or past tense). See also Dahl (1984, 112–113)
Metadiscursive tokens were excluded in the statistical analysis of the variation between *be going to* and other future temporal reference forms because their small number and unique patterning. It is worth keeping in mind that these types of tokens would otherwise be coded as proximate because the event time is immediately following speech time. It may prove fruitful to re-investigate the data from the above communities to determine whether proximity effects are still operational once these highly conservative proximate contexts are set aside.

Metadiscourse  And I'll tell you, that was scary for Sydney. (Female born 1961, PI)
Metadiscourse  If you were, and I'm going to go there, if you were a lesbian, it didn't matter. (Female born 1966, PI)
Metadiscourse  There was one little fellow. I won't mention names. He was a little Newfoundlander and he was about as sharp operating that furnace as you want to get. (Male born 1917, PI)
Proximate  I dread going to work tonight because it's going to be pure hell. (Male born 1954, PI)
Distal  I'm not going to come back here next Tuesday night and tell you anything different. (Male born 1926, SW)
Unspecified  So I said, "I'm going to say something that only a Cape Breton would get." (Female born 1956, PI)

As with Guysborough Village there is a clear correlation in Cape Breton between proximate contexts and the use of *be going to* in the combined dataset. When the reference time is specified as 24 hours or less from the speech time *be going to* is used at at rate of 55 percent, but when the reference time is specified as more than 24 hours from speech time *be going to* is used at a rate of 37 percent. When the reference time is not specified any more specifically than simply the future, *be going to* is used 43 percent of the time. This rate is between those of specifically proximate and specifically distal contexts. While the unspecified proximity rate is closer to that of distal contexts, and like distal contexts below 50 percent, it is nearly identical with the overall rate of 42 percent. Whether the difference between proximate, distal, and unspecified contexts is significant above and beyond the random effects of speaker and when considered alongside other potentially constraining factors will be tested in Section 9.6.4.7. Further, if there is a significant contrast, whether that contrast is three-way or two-way, and if two-way, whether unspecified contexts are statistically distinct from proximate or distal contexts, will be
important in assessing the similarity of the Cape Breton future temporal reference system to the system in the British Isles described by Tagliamonte, Durham, and Smith (2014), where unspecified contexts and proximate contexts are not statistically different from one another and favour be going to relative to distal contexts.

The “odd duck” among proximity factors is metadiscourse; in other words, contexts in which the reference time is the speech time and the reference event is the adjacent utterance itself. There were a total of 262 metadiscursive tokens across the Storyteller, Steelworker, and Post-Industrial corpora. The majority (66 percent) are the frame I_____ tell you…, in which, will occurs 88 percent of the time and be going to occurs only 12 percent of the time. This, of course, is a much higher rate of will and a much lower rate of be going to than the overall frequencies in any of the corpora. Of the remaining 90 tokens, only 17 percent occur with be going to, again much lower than the overall rate for any of the three corpora. Just as with deontic modality (Section 8.6.1) metadiscursive contexts appear to be especially resistant to the incoming form, which in this case is be going to. For future temporal reference, this phenomenon is best exemplified by I’ll tell you…, which is likely more “formulaic” and less compositional in its meaning than the other expressions (e.g., I won’t mention names…, or I’ll be honest with you…). In other words, I’ll tell you… is more firmly entrenched as a discourse marker, ergo more resistant to innovation. When all of the metadiscursive tokens were considered together the overall rate of be going to was only 13 percent. The rate of be going to in metadiscursive contexts also declined in real time, as the overall use of be going to increased, suggesting that these types of constructions are becoming more specialized for will. In the Storyteller corpus, where the overall rate of be going to is 27 percent, the rate of be going to in metadiscursive contexts is 26 percent percent; in the Steelworker corpus, where the overall rate of be going to is 44 percent, the rate in metadiscursive contexts is 10 percent; and in the Post-Industrial corpus, where the overall rate of be going to is 42 percent, the rate in metadiscursive contexts is 8 percent. The same pattern also exists in apparent time. A conditional inference recursive tree analysis of the combined dataset with metadiscursive tokens included identified 1908 and 1948 as shock points in apparent time (see Figure 9.6). Among speakers born before 1908, the overall rate of use of be going to is 25 percent, while the rate of be going to in metadiscursive contexts is 22 percent. For speakers born between 1908 and 1948 the overall rate of be going to is 42 percent and the rate of be going to in metadiscursive contexts is 11 percent. Finally, for speakers born after 1948, the overall rate for for be going to is 48 percent, yet the rate of be going to in metadiscursive contexts is 10 percent.

Of note, the rate of be going to with metadiscursive I_____ tell you… drops from 18 percent to 9 percent to 8 percent in apparent time for the three age categories mentioned above. All other metadiscursive contexts decline from 50 percent to 13 percent then remain 13 percent. This indicates that while I’ll tell you… may lead in specialization with will in apparent time, the high rates of I’ll tell you… relative to other metadiscursive expressions is not solely responsible for the overall trend of be going to decline in metadiscursive contexts. The rate of be going to for tell in non-metadiscursive contexts, as in (83) — of which there is much less data — is 22 percent. This is an important observation because it provides evidence that the discursive context was likely the root of the collocation effects for I’ll tell you… observed in Québec City and Montréal English by Torres Cacoullos and Walker (2009). The authors do propose I’ll tell you… could be acting as a discourse marker in their data, but stop short of probing this explanation any further. This corroborating evidence from the Cape Breton data indicates they were likely correct in their observation. By motivating the genesis of this collocation effect — via showing I’ll tell you… is functioning as a metadiscursive perlocutionary discourse marker and it is its status as such that resists innovation — Torres Cacoullos and Walker’s (2009) findings can

---

17. I’ll tell you is also prototypical as a perlocutionary marker, which are expressions that make overt commentary on an adjacent utterance. The denotational meaning of I will tell you… is ‘I declare this utterance’.
be reconciled with Tagliamonte and D’Arcy’s (2009) astute observation that most grammatical contexts do not show the extreme imbalance that would result from collocation effects alone generating constraints on will and be going to variation.

(83) If they’re looking to shine bright in the eyes of management, they’re going to tell management what management wants to hear. (Male born 1959, PI)

Figure 9.6: Conditional inference recursive partitioning tree — be going to vs. will in Cape Breton English. Parameters included: year of birth and sex.

9.6.4.3 Verb Type

A total of 432 different verbs occurred with future will and be going to in the combined dataset. Table 9.3 lists those verbs that occurred more than 47 times, and lists the verbs that occurred only once. Given that Torres Cacoullos and Walker (2009) find putative collocation effects for some verbs, and that different verbs simply occur with each variant at wildly different frequencies (e.g., lose 67 percent vs. find 30 percent be going to be going to), specific verb was included as a random effect in the regression analysis in Section 9.6.4.7. Here, lexical be and auxiliary be were labeled separately, as were stative possessive have and deontic modality have to.

Both its overall frequency of occurrence in the dataset and the exceedingly high rate of will, make the verb notice somewhat unusual. For example, fifteen of the most frequent verbs listed in Table 9.3 are among the top 100 most common words in the Corpus of Contemporary American English (COCA; Davis 2008). Of the remaining,

18. 47 tokens represents about 1 percent of total dataset
19. Auxiliary and lexical be are considered as one verb, as are possessive and deontic have in the COCA.
Table 9.3: Distribution of expressions of future temporal reference in Cape Breton English — variant by verb.

<table>
<thead>
<tr>
<th>Verb</th>
<th>will</th>
<th>be going to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>be + adjective</td>
<td>236</td>
<td>47</td>
</tr>
<tr>
<td>get</td>
<td>146</td>
<td>48</td>
</tr>
<tr>
<td>go</td>
<td>157</td>
<td>61</td>
</tr>
<tr>
<td>do</td>
<td>68</td>
<td>31</td>
</tr>
<tr>
<td>tell</td>
<td>187</td>
<td>86</td>
</tr>
<tr>
<td>have</td>
<td>96</td>
<td>55</td>
</tr>
<tr>
<td>have to</td>
<td>84</td>
<td>57</td>
</tr>
<tr>
<td>see</td>
<td>128</td>
<td>89</td>
</tr>
<tr>
<td>take</td>
<td>88</td>
<td>65</td>
</tr>
<tr>
<td>be + past participle</td>
<td>55</td>
<td>47</td>
</tr>
<tr>
<td>say</td>
<td>92</td>
<td>79</td>
</tr>
<tr>
<td>be + present participle</td>
<td>72</td>
<td>69</td>
</tr>
<tr>
<td>work</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>put</td>
<td>64</td>
<td>67</td>
</tr>
<tr>
<td>make</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>give</td>
<td>65</td>
<td>77</td>
</tr>
<tr>
<td>come</td>
<td>53</td>
<td>65</td>
</tr>
<tr>
<td>notice</td>
<td>55</td>
<td>98</td>
</tr>
<tr>
<td>happen</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>

The following verbs occurred only once. Underlined verbs occurred with will: accept, allow, amalgamate, analyze, answer, appreciate, attach, attempt, babbitt, back, backfire, bang, bar, bat, belong, benefit, bid, bite, bleed, bless, boil, borrow, boss, breathe, bulldoze, bump, button, bypass, cancel, chip, clean, clear, cooperate, compensate, comprise, confront, construct, contract, control, convert, correct, count, create, creep, crunch, curse, damage, dance, defy, de-gas, deliver, demand, demolish, describe, dial, differentiate, disappear, discard, dispute, doctor, dress, drill, duck, enter, erect, evolve, exhaust, explode, expose, face, feed, finance, flip, flow, fluctuate, flush, fly, fool, genuflect, graduate, grind, grow, guarantee, guess, hammer, harden, heat, identify, intercede, invite, jeopardize, jive, justify, knock, lance, lean, level, lifeguard, load, lunch, man, manage, mar, march, melt, milk, modernize, monitor, mud, negotiate, note, object, oblige, offend, overflow, pace, pen, perform, pile, pin, pipe, please, poison, poke, predict, press, prove, provide, purify, radio, raise, ramp, re-assess, re-heat, react, realize, receive, recognize, refuse, relocate, rent, repair, rephrase, require, reserve, resolve, retrain, rig, ring, rise, roam, scare, second, select, shape, shatter, shim, shit, shoot, shrink, shunt, skip, slip, slow, smell, solve, specialize, spit, splash, spoil, spread, sprinkle, squeal, stall, steal, step, stir, strike, strive, stump, submit, succeed, sue, support, swear, switch, tie, tighten, trick, trim, understand, up, update, vibrate, wake, weigh, welcome, win, wind, wipe, zip.

tell is ranked 103rd, work is 117th, and happen is 217th. The verb notice on the other hand is 1048th. This anomaly is an idiosyncrasy of the Steelworker corpus. A large portion of the discussion that occurs in the corpus involves the interviewee explaining to the interviewer details of the steelmaking process pictured in historical photographs. Frequently the interviewees use You’ll notice... as a deictic expression when pointing out details in
the photograph. Fifty-two of the 56 notice tokens are this expression used deictically from the Steelworker corpus. As with metadiscursive perlocutionary markers, the reference time for these deictic constructions is speech time, and while the reference event is not the utterance itself, the reference event is part of the interaction between the speaker and the interlocutor(s). You’ll notice... is ungrammatical as an utterance on its own; it is instead used to draw attention to the content of a declaration. In (84) below the speaker uses you’ll notice to draw attention to a fact about the image that the speaker wants the interview to see, and in doing so to know. You’ll notice is just one of several discourse strategies (e.g., discourse marker now, deictic this here, and tag questions) the speaker uses to organize the short utterance, to connect the content of his utterance to the photograph being examined and to connect the content of his utterance to the understanding of the interviewer. In this sense, you’ll notice’s meaning is pragmatic, rather than denotational, and thus is likely constructed and understood with less focus on its compositional syntactic/semantic meaning. You’ll notice most likely exists in the grammar as a “chunk,” like I must say or I’ll tell you, that resists innovations. As with metadiscursive contexts, tokens with the verb notice were laid aside and excluded from the regression analysis.

\[(84)\] Now this here is number two shop, see the big cranes you got? Now you’ll notice it’s much cleaner. Much more modern, isn’t it? Compared to the ones we were looking at. (Male born 1923, SW)

Given that over 200 verbs occurred with only one token in Table 9.3, there are many instances of categoricity by verb in the data. Verb was therefore considered as a random effect in order to take this uneven distribution into account.

The patterning of verb groups; however, was considered as a fixed effect. As described in Section 9.3, the will future grammaticalized from a modal verb denoting desire or volition. The be going to future grammaticalized from a lexical verb denoting movement towards a goal. For this reason there may be a higher instance of will with stative verbs, and a higher rate of be going to with dynamic verbs. However, Poplack and Tagliamonte (2000, 335) argue that the use of be going to with another dynamic verb instead implies semantic bleaching or desemanticization of its original lexical content and is indicative of a later stage of grammaticalization. They support this argument by pointing out that in non-mainstream (e.g., isolated, conservative) varieties (Samaná, North Preston, Guysborough Enclave, and Guysborough Village) verbs of motion significantly favour the use of will relative to be going to, while in mainstream varieties (e.g., Ottawa) — ostensibly further along the cline of grammaticalization — this constraint has neutralized. Tagliamonte, Durham, and Smith (2014) do not test verb type in their analysis of British dialect data. Torres Cacoullos and Walker (2009) do test several verb categories, which include differentiating telic and atelic verbs, as well as transitive and intransitive verbs; however, the relative rates of be going to across groups was not statistically significantly different.

For the present study verbs were categorized based on aktionsart following Vendler’s (1957) classification. Non-dynamic, e.g., stative verbs, like be, stay, and live, were classified as states; non-punctual, durative dynamic verbs like work and dance were categorized as activities; punctual, non-durative dynamic verbs like retire and arrive were categorized as achievements; and punctual, durative dynamic verbs like harden and make were categorized as accomplishments.

\[(85)\] State I’ll be 92 my birthday. (Male born 1880, ST)

Activity And it’s like, “Okay, go over and ask her if she’ll dance with me.” (Male born 1958, PI)

Achievement It is a relatively young guy, yeah, like the kind that you think is going to die in a hole some day in a gutter. (Male born 1991, PI)

Accomplishment I told him, “Yes, I’m going to write that story.” (Male born 1950, SW)

Distributionally there was little difference between the four categories, ranging from 37 to 51 percent be going to use. While it is true that be going to is used more frequently with non-dynamic verbs (e.g., states),
potentially aligning Cape Breton English with the other conservative Nova Scotia varieties described by Poplack and Tagliamonte (2000), whether the distinction is statistically significant will be determined by the regression analysis.

Figure 9.7: Distribution of *be going to* by proximity in Cape Breton English.

### 9.6.4.4 Clause Type

Across studies of the grammaticalization of *be going to*, a consistent effect of clause type is attested. *Be going to* is favoured in subordinate clauses, while disfavoured in main clauses (Royster and Steadman 1923/1968; Poplack and Tagliamonte 2000; Torres Cacoullos and Walker 2009; Tagliamonte, Durham, and Smith 2014). As Tagliamonte, Durham, and Smith (2014, 91) point out, this may be because subordinate clauses were the entry point for *be going to* to grammaticalize for expressing future.

(86) **Apodosis** He said, "If you give me two dollars, I’ll give you all the rum you want." (Male born 1890)

Other Main Clause **Oh** yeah. *I’ll never forget September, 11th.* (Female born 1991)

Other I think [that] they’re going to get up there sooner or later. (Male born 1955)

In both North Preston and Guysborough Enclave, Nova Scotia, Poplack and Tagliamonte (2000) find that clause type is a significant predictor, but that the distinction is not significant in either Guysborough Village or Ottawa. Torres Cacoullos and Walker (2009) find a three-way distinction, whereby subordinate clauses favour *be going to* (62 percent), while main clauses, and even more so, apodoses favour *will* (48 percent and 44 percent *be going to* respectively). Apodoses are the main (consequent) clauses of a conditional sentences; in other words, they contain futures that depend on the satisfaction of some condition. The authors hypothesize that if one of the future forms conveyed more uncertainty (a claim often made by prescriptivists aiming to semantically distinguish between future variants), it would be more common in apodoses. While the factor group of clause type was significant in Torres Cacoullos and Walker’s (2009) analysis, the authors did not test the internal contrasts among factor groups with more than two levels. Subordinate clauses, with 62 percent *be going to* use are clearly distinct from the two groups of main clauses, but whether apodoses at 44 percent and other main
clauses at 48 percent are distinct is unclear. Tagliamonte, Durham, and Smith (2014) find a significant two-way distinction between be going to favouring subordinate clauses and disfavouring main clauses in their combined British dialect dataset, and in each individual community. The authors originally tested the three-way distinction presented by Torres Cacoullos and Walker (2009), with apodoses considered separately from other main clauses, but found that the only relevant distinction was between subordinate and main clauses (Tagliamonte, Durham, and Smith 2014, 93, Footnote 26). The authors also found that this constraint was less strong among younger speakers and in less peripheral communities.

Figure 9.8 shows that an actual three-way distributional contrast exists in the Cape Breton data. Be going to occurs 29 percent of the time in apodoses, 40 percent of the time in other matrix clauses, and 61 percent of the time in subordinate clauses — though whether these differences are other than chance variation must be determined using statistical tests. While a three-way distributional contrast may exist, it is also important to remember that this is the conflation of two independent two-way contrasts: principle main clauses versus subordinate clauses, and apodosis versus all other clauses. It is possible for an apodosis that is also a subordinate clause, as in (87) below. These types of constructions are very rare in the dataset, and so the three-way factor group matching Torres Cacoullos and Walker’s (2009) was employed.

(87) a. Now, it’s touchy in the sense that if you’re going to put it down and you put it down too hard, you feel like you’re going to fall out the window! (Male born 1947, SW)
b. I’m not the guy that if you tell him to do something he’ll go out on his own and do it. (Male born 1948, SW)

9.6.4.5 Temporal Adverbs

Torres Cacoullos and Walker (2009) find that non-specific temporal adverbs (or adverbial phrases), like soon, later, and someday favour will relative to specific temporal adverbs (or adverbial phrases) like now, tomorrow, or next Tuesday at 11 am. The authors posit that the association may result from will connection to uncertainty.
That being said, the authors also recognized that the significance of this factor group may be driven by the collocation effects of *will never*. In their dataset *will* occurs in 89 percent of their *never* tokens. As noted above, this pattern also occurs in the Cape Breton data, where *never* futures occurs with *will* 90 percent of the time.

(88)  Specific  *Tomorrow, we’ll show him what he’s going to get.* (Male born 1927)
       Non-specific  *I’ll have to drag you there sometime.* (Female born 1990)
       No adverb  *I don’t know how you’re going to get them out yet, pardon the expression.* (Male born 1950)
       Never  *The next thing he promised, he said, “I’ll never drink another drop of booze.”*

There is clearly some kind of collocation effect occurring with *never* and *will*. Unlike perlocutionary markers, or deictic *You’ll notice*, for which the collocation effects likely result from these types of expressions resisting innovations — a phenomena observed for both future temporal reference and deontic modality — the cause of the high rate of *will* with *never* appears to be a mystery. It is especially mysterious given that *will* is associated with uncertainty, and *never* specifically denotes not just a punctual but a durative negation — in other words, certainty in negation. While historically, and in contemporary non-standard varieties *never* is used for simple negation (e.g., Cheshire 1999, 32; Clarke 2010, 98), since the mid-18th century the prescribed usage for negation is with universal temporal quantification. In this sense, it aligns with *will*’s slight favouring in distal contexts. Torres Cacoullos and Walker (2009) identify this trend in their data, and suggest that it arises because all future constructions are on a continuum of lexical specificity and productivity, of which *X’ll never + Verb* and *I’ll tell you* occupy an extreme end. They propose that fixed collocations reflect not just older meanings (à la Bybee and Pagliuca 1987) but also distribution patterns, as they have conventionalized. The authors; however, do not identify the original distribution pattern or meaning that lead to strong collocation effect of *never* and *will*, nor do they explain why *X’ll never + Verb* conventionalized while other constructions did not.

Based on Cheshire’s (1999) analysis of *never*-use from an interactional perspective, I suggest the same phenomena of perlocutionary/discourse markers resisting innovation may be at work. Cheshire (1999, 42) points
out that never “frequently occurs in a cluster of addressee-oriented forms at the beginning of a narrative, when the speaker needs to secure the interest of the interlocutor in order to keep the floor for an extended turn.” She explicitly references I’ll never forget... as the prototypical expression for this perlocutionary/discourse function, though other expressions like You’ll never guess... serve the same function. In the Cape Breton data, of the 94 tokens with never negation only four verbs occur more than three times each: lexical be (11 tokens), get (8 tokens) and see (9 tokens), the first, second and eighth most common verbs in the dataset, as well as forget, with 29 tokens, representing about one-third of the data.20 One hundred percent of the never + forget tokens occur with will. Furthermore, 100 percent of the negative forget tokens occur with never. Never forget looks like another metadiscursive “chunk” where innovation may be inhibited or decelerated.

Removing the never forget tokens from the never-negated tokens does not automatically solve the mystery of the never + will collocation effect. Will still occurs 88 percent of the time in the 65 remaining never negated tokens. Looking at the eight tokens of never + be going to perhaps helps unravel the mystery though. Seven of the tokens come from speakers from the Steelworker corpus, all speakers born before 1948. The remaining token comes from a male born in 1878 in the Storyteller corpus talking about WWI, (89). This means that no never + be going to tokens occur in the most recently collected data, or among the youngest age cohort in apparent time as identified by the conditional recursive tree analysis. This is consistent with the proposal above that as be going to continues to increase in frequency over apparent/real time, the specialization of will for specific perlocutionary functions becomes more entrenched. This process of specialization, however, need not be limited to metadiscursive functions. The specialization of will may expand to include all never constructions, with never forget as its model.

(89) I told Steve, “I’m not a miner,” and he said, “Shut up. You want to see France, don’t you? First thing we know, the war will be over, and the 64th is never going to get there.” (Male born 1978 ST)

Given the pattern of never negation future constructions, and the fact that the never tokens have the potential to skew both the adverb type and sentence type factor groups, these tokens were excluded from the regression analysis in Section 9.6.4.7. Furthermore, as adverbs signal proximity, all tokens with no adverbs are unspecified for proximity. This means these two factor groups overlap. Beyond that, the distinction between specific adverbs, non-specific adverbs, and no adverbs is incredibly small. For this reason, beyond the identification of never as an outlier temporal adverb, this study will not test the specific semantic qualities of temporal adverbs.

9.6.4.6 Subject Type

Two qualities of the subject have been shown to constrain the variation between be going to and will: subject person and subject animacy. Below, the observed pattern for the grammatical person of the subject will be detailed, followed by the observed pattern for subject animacy. Finally, a scheme for coding tokens based on these two qualities will be outlined.

The first trend observed across studies is that first person subjects favour will, while second and third person subjects favour be going to (Wekker 1976; Poplack and Tagliamonte 2000; Torres Cacoullos and Walker 2009).

(90)

First Person It was just the fact that this rivalry exists for this long, and we’re going to keep it up. (Female born 1966, PI)
Second Person You’re not going to get them to show you how to do it? (Female born 1940, PI)
Generic Pronoun If you start talking about someone behind their back, then you’re going to have problems probably. (Male born 1991, PI)

20. You’ll never guess... does not occur in the Cape Breton data.
Third Person Pronoun Now, he, you know, if he knows the regular name, he will use it. But Glace Bay, and the coal mine, was terrible, terrible for nicknames. (Female born 1966, PI)

Third Person Noun Phrase Your car is not going to work very well, it’s not going to be efficient. If it’s not efficient, you can’t afford to run it. (Male born 1945, PI)

Tagliamonte (2002, 750) points out that will is usually associated with volition (see also, Quirk et al. 1985, 213–214, 238–229; Tagliamonte, Durham, and Smith 2014), and first person subjects are more likely to show volition. Of note is also the prescriptive tradition, described in Section 9.3, that shall is the correct future form for first person singular subjects, while will is the correct for other subjects.

This trend was significant in both North Preston and Ottawa, but not in Guysborough Village (Poplack and Tagliamonte 1991). Likewise it was significant, and in the same direction, in Québec City and Montréal (Torres Cacoullos and Walker 2009). In Tagliamonte, Durham, and Smith’s (2014) British dialect data, the constraint is clearly active among younger speakers and in less peripheral varieties, but it is unclear whether it is operational among the older, peripheral speakers — which includes the oldest speakers from Scotland and Northern Ireland.

A second trend that has been observed across studies, is that human animate subjects favour be going to. This may be because be going to originates from a verb of movement towards a goal, and that movement requires an animate subject. As it grammaticalizes, and its meaning desemanticizes and bleaches, be going to expands to be used with non-human animate and then finally inanimate subjects. Alternatively, be going to may have been available for all subject types from its actuation as a future temporal reference expression and has grown in parallel at a constant rate in these contexts, but its initial frequency with human subjects (those capable of movement towards a goal) was greater than with other subjects.

(91) Animate Dorothy is going to be singing Somewhere Over the Rainbow. (Male born 1991, PI)  
Animate Of course, a bear’s is, too - but no bear is going to walk on his two rear legs four or five miles. (Male born 1945, ST)  
Animate I’d be nervous taking my dog on the ferry. You never know how they’ll react, right? (Male born 1987 PI)  
Animate They’re going to be like “Well, we got to get the fuck rid of the stop sign.” (Male born 1989, PI)  
Animate You’re going to have inspectors that will say, “Hold it, you don’t have that thing hot enough, look, she won’t burn that. That will go out in the atmosphere amongst the people.” (Male born 1942, SW).  
Non-animate You’re going to have inspectors that will say, “Hold it, you don’t have that thing hot enough, look, she won’t burn that. That will go out in the atmosphere amongst the people.” (Male born 1942, SW).  
Non-animate ’Cause I know once I start thinking about it, my accent will get stronger. (Female born 1990, PI)  
Non-animate It is hot up in the compartments. You can still work in there. It won’t burn you or nothing. (Male born 1944, SW)

In the Cape Breton data both humans and animate non-humans, like bears and dogs, were coded as animate. This includes all first person and second person subjects, including general you, as in the fifth example in (91). General they, as in the fourth example, was also coded as animate. Organizations or groups that may exert volition, but that are nonetheless non-human, like the company or the government, were coded as inanimate; as were existential subjects, like existential it and there. Finally, non-animate referred to using he or she, as in the sixth example in (91) (see Section 3.2), were nonetheless coded as inanimates. Given that subject number and subject animacy overlap in the data (all first and second person subjects are animate), these factors were combined and the three-way distinction between first person, second/third person animate, and third person non-animate was tested.

In the Cape Breton data first person subjects have the highest rate of will, as expected, but it is non-animate subjects that lead in the use of be going to, defying expectations. A careful examination of the combined effects of multiple constraints over time, and over corpora may explain this unexpected pattern.
9.6.4.7 Logistic Regression Analysis

The first step in assessing the combined effect of multiple parameters on the variation between be going to and will in the spoken data from the Storyteller, Steelworker, and Post-Industrial corpora is to determine how to assess change over time. To do this a conditional inference recursive partitioning tree (Section 5.1, p. 5.1) was created, testing the combined effect of corpus, sex, and year of birth. The conditional inference tree, presented as Figure 9.11, is different from the conditional inference tree presented as Figure 9.6 because metadiscursive and deictic contexts and expressions with never have been excluded, and corpus has been included as a test parameter. Figure 9.11 shows that the data from the Storyteller corpus is significantly different (p < 0.001) than the data from the other two corpora. Further, within the Storyteller corpus neither year of birth nor sex are significant predictors. In the regression analysis the data from the Storyteller corpus is analyzed separately, representing the oldest data in both real and apparent time. The Steelworker and Post-Industrial corpora are combined, because the conditional inference tree analysis reveals that there is not a significant difference between the two corpora among women or among male speakers born after 1948. Therefore, the combined Steelworker-Post-Industrial data will be partitioned by year of birth, with 1948 as the dividing year.

That men are ostensibly leading the change towards be going to — something that must be confirmed after other parameters are taken into account — suggests that the variant be going to is judged as non-standard or non-prestigious (Trudgill 1972). This aligns with traditional prescriptive accounts of be going to, and especially gonna, being informal (Wekker 1976; Mair 1997, 1541; Berglund 2005; Szmrecsanyi 2003, 303). It further aligns with specific patterns in Nova Scotia, whereby marginalized sub-communities (e.g., African–Nova Scotians) and traditional communities without primacy (e.g., Loyalists) have relatively higher rates of be going to.

Partitioning the data based on the results presented in Figure 9.11 provides three distinct groups that can also roughly represent periods of linguistic development. The Storyteller corpus, collected in the 1970s and early 1980s, is the oldest real-time data, and the data that extends the furthest back into apparent time. It also
Figure 9.11: Conditional inference recursive partitioning tree — be going to vs. will in Cape Breton English. Parameters included: year of birth, sex, and corpus.

The data from these two corpora come from urban-industrial rather than rural areas, and include a mix of working class and professional speakers. The data from the third group is also from both the Steelworker and Post-Industrial corpora and represents the “newest” data. Of note is that there are no speakers in the Steelworker corpus born after the 1960, so any age effects found in this third group may represent real time in addition to apparent time change.

---

21. See Section 4 for a discussion of the descriptors “rural” and “urban-industrial” in a Cape Breton context.
Table 9.4: Mixed effects logistic regression of the contribution of factors to the probability of *be going to* for expressing future temporal reference across corpora and age groups in Cape Breton English.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>aic</em></td>
<td>512</td>
<td>3,102</td>
<td>1,457</td>
</tr>
<tr>
<td><em>input</em></td>
<td>.46</td>
<td>.64</td>
<td>.67</td>
</tr>
<tr>
<td>total N</td>
<td>471</td>
<td>2,722</td>
<td>1,144</td>
</tr>
</tbody>
</table>

**Fixed Effects**

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question</strong></td>
<td>.75</td>
<td>65</td>
<td>.86</td>
<td>94</td>
<td>.91</td>
<td>95</td>
</tr>
<tr>
<td><strong>Negative Statement</strong></td>
<td>.45</td>
<td>36</td>
<td>.32</td>
<td>51</td>
<td>.24</td>
<td>48</td>
</tr>
<tr>
<td><strong>Affirmative Statement</strong></td>
<td>.29</td>
<td>25</td>
<td>.26</td>
<td>41</td>
<td>.24</td>
<td>49</td>
</tr>
<tr>
<td>range</td>
<td>46</td>
<td>60</td>
<td>60</td>
<td></td>
<td>67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clause Type</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Other Clause</strong></td>
<td>.64</td>
<td>38</td>
<td>.68</td>
<td>63</td>
<td>.69</td>
<td>69</td>
</tr>
<tr>
<td><strong>Other Main Clause</strong></td>
<td>.43</td>
<td>26</td>
<td>.46</td>
<td>39</td>
<td>.51</td>
<td>42</td>
</tr>
<tr>
<td><strong>Apodosis</strong></td>
<td>.44</td>
<td>22</td>
<td>.35</td>
<td>31</td>
<td>.31</td>
<td>39</td>
</tr>
<tr>
<td>range</td>
<td>21</td>
<td>33</td>
<td>33</td>
<td></td>
<td>33</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject Person/Animacy</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Animate</td>
<td>[.43] 15</td>
<td>84</td>
<td>.64</td>
<td>60</td>
<td>.62</td>
<td>64</td>
</tr>
<tr>
<td>Other Animate</td>
<td>[.61] 30</td>
<td>229</td>
<td>.53</td>
<td>43</td>
<td>.51</td>
<td>50</td>
</tr>
<tr>
<td>First Person</td>
<td>[.40] 30</td>
<td>158</td>
<td>.33</td>
<td>27</td>
<td>.37</td>
<td>38</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td>31</td>
<td></td>
<td></td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proximity</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximate</td>
<td>[.61] 37</td>
<td>19</td>
<td>.65</td>
<td>60</td>
<td>.61</td>
<td>60</td>
</tr>
<tr>
<td>Unspecified</td>
<td>[.43] 27</td>
<td>421</td>
<td>.43</td>
<td>44</td>
<td>2,503</td>
<td>.52</td>
</tr>
<tr>
<td>Distal</td>
<td>[.46] 32</td>
<td>31</td>
<td>.42</td>
<td>38</td>
<td>128</td>
<td>.41</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td>23</td>
<td></td>
<td></td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb Type</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
<th>FW %</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>[.58] 30</td>
<td>177</td>
<td>[.48] 43</td>
<td>1,105</td>
<td>[.49] 49</td>
<td>490</td>
</tr>
<tr>
<td>Accomplishment</td>
<td>[.64] 32</td>
<td>116</td>
<td>[.52] 40</td>
<td>915</td>
<td>[.49] 47</td>
<td>252</td>
</tr>
<tr>
<td>Achievement</td>
<td>[.33] 21</td>
<td>71</td>
<td>[.48] 38</td>
<td>370</td>
<td>[.47] 45</td>
<td>121</td>
</tr>
<tr>
<td>range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
Table 9.4 – continued from previous page

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Year of Birth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Continuous factor</em></td>
<td>[—]</td>
<td>[—]</td>
<td>+0.0175 log-odds</td>
</tr>
<tr>
<td><strong>Random Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Variance N</td>
<td>Variance N</td>
<td>Variance N</td>
</tr>
<tr>
<td></td>
<td>1.40 74</td>
<td>0.483 140</td>
<td>0.267 78</td>
</tr>
<tr>
<td>Verb</td>
<td>1.00 117</td>
<td>0.822 314</td>
<td>0.090 219</td>
</tr>
</tbody>
</table>

Sum contrast coding.
- Aic, input, and factor weights for significant factor groups derived from most parsimonious model. Factor weights for non-significant factor groups derived from full model.
- Factor weights converted from log-odds. \( p = \frac{e^x}{1+e^x} \).
- Factor weights with non-significant contrasts indicated with [ ].
- Correlation of fixed effects: Storyteller \( r < |0.02| \), Steelworker/Post-Industrial - 1902–1948 \( r < |0.065| \), Steelworker/Post-Industrial - 1949–1999 \( r < |0.052| \).
- Additional treatment contrast coding used to test orthogonal contrasts for parameters (factor groups) with more than two levels (factors) using most parsimonious model.
- Significant contrasts indicated with horizontal lines.
- † Note order from most favouring to least favouring is Third Person, Animate >Other

Table 9.4 lays out the results of a mixed-effects logistic regression analysis using the lm4 package in R (Bates, Maechler, et al. 2015; R Core Team 2015). This model tests the fixed effects of sentence type, clause type, subject (grammatical person plus animacy), proximity, verb type, sex and centred year of birth, taking into account the potential random effects of individual speaker and verb. As with previous chapters the results are organized in a format that is consistent with variationist literature, with log-odd likelihoods converted into factor weight probabilities.22 Square brackets indicate that distributional differences between parameter levels (e.g., factors) are not significant when other parameters (e.g., factor groups) are taken into account. This was determined using a Wald \( \chi^2 \) test on the full model (see Footnote 6 on p. 6). For parameters with more than two levels, additional statistical tests were performed to determine which internal contrasts were significant. These contrasts are indicated with horizontal lines.

The first thing notable about the data presented in Table 9.4 is the overall input of the model. The overall likelihood of *be going to* increases from .44 in the Storyteller data, to .64 in the older mixed data, and .67 younger mixed data. This indicates that there is an increase in the use of *be going to* occurring in the community, but perhaps the change may be slowing down. This is also the logical conclusion drawn from the distribution of forms across corpora presented in Table 9.2. However, the year of birth parameter for the youngest cohort — Steelworker/Post-Industrial speakers born after in the second half of the 20th century — does show a significant positive effect, indicating that as year of birth increases so to does the likelihood of *be going to*. Taken together with the results of the conditional inference recursive partitioning trees (Figure 9.11), it appears that this increase

---

Figure 9.12: Distribution of *be going to* by sex in Cape Breton English. Speakers grouped by decade of birth. Due to the small number of tokens (< 5) from female speakers in the 1930s (from 1937) and 1970s (from 1974) this data was added to the 1940s and 1960s cohorts respectively. There were no female tokens from the 1880s. Also, due to the small number of male tokens from the 1970s (< 5), from 1973) these tokens were added to the 1960s cohort.

is specifically among males in the *Post-Industrial* corpus. Figure 9.12 shows the rise of *be going to* across the three datasets by sex. Across most of the 20th century the rise of *be going to* is lead by males; however, the difference between males and females increases among the youngest speakers.

The dip in *be going to* use among the speakers born in the 1990s, which is clearly visible in 9.12 is likely an adolescent peak rather than retrograde change, given that speakers from this decade were younger than 20 when interviewed. The same adolescent peak is observed for quotative *be like* in Figure 6.6, and was found for *be going to* by Denis and Tagliamonte (2017). Even though here *be going to* is a male-lead innovation, adolescent peaks for male speakers are unexpected; adolescent incrementation (of which adolescent peaks are a consequence) is theorized to only occur for females, leading to sex differentiation in variant use (Labov 2001, though see Tagliamonte and D’Arcy 2009 for a discussion of sex differentiation and male adolescent peaks).

Looking to the linguistic constraints, across all three speaker groups the effect of sentence type is significant and has the greatest magnitude of effect. For each of the three groups interrogative contexts favour *be going to*, while both negative and affirmative declarative contexts lag behind. Further, in each group there is only a statistically significant two-way contrast between questions and statements; the difference between negative and affirmative contexts is not significant. While the difference between negative and affirmative statements is non-significant, it is worth noting that the difference in rates of *be going to* use between these two categories is shrinking with time. In the *Storyteller* data *be going to* occurs with negative statements 36 percent of the time versus 25 percent of the time for affirmative statements. Those rates increase to 51 percent and 41 percent in the older partition of the mixed data and 48 percent and 49 percent in the younger partition of the mixed data, suggesting perhaps that the two environments, while increasing their *be going to* use, are not doing so at the same rate over time (anathema to the Constant Rate Effect, Kroch 1989, and the findings presented by Denis
and Tagliamonte 2017). Figure 9.13 shows the predicted probability by birth year and sentence type based on a mixed effects logistic regression model constructed identically to Table 9.4, but without the data partitioned. Each point in Figure 9.13 represents one token, plotted by year of birth and its predicted probability given its particular combination of parameter levels settings. Larger points represent multiple tokens with the same combination of parameter level settings. Orange circles are affirmative statement tokens, green triangles are negative statement tokens, and blue squares are question tokens. Each line is a binomial curve fitted to the predicted probabilities of each sentence type. The shaded bands represent the 95 percent confidence interval. These bands increase in size at either end because there are fewer speakers born before 1900 or after 1995. This figure shows that across apparent time, in spite of other constraining factors, the probability of *be going to* with questions is very high, but with a flat trajectory. This is expected for changes that have gone to completion in a given context. The probabilities of *be going to* in affirmative and negative statements grow in parallel across the 20th century, with negative statements consistently more probable than affirmative statements, but each similar enough that their confidence intervals overlap across much of their apparent time trajectories. Figure 9.13 indicates, contrary to the conclusions drawn from the usage percentages presented in Table 9.4, that *be going to* is in fact increasing at a constant rate in the contexts where it has not already reached completion (aligning with Kroch 1989, and the findings presented by Denis and Tagliamonte 2017).

![Figure 9.13: Predicted probability of *be going to* by sentence type and year of birth in Cape Breton English. Probability calculated based on a mixed effects logistic regression model of the full dataset with the fixed effects of sentence type, clause type, subject (grammatical person plus animacy), proximity, verb type, sex and centred year of birth, and the random effects of individual speaker and verb.](image)

Turning to clause type, subordinate clauses have the highest rate of *be going to* use across the three partitions, and apodoses have the lowest (and therefore the highest rate of *will*), it is only among the speakers in the mixed data that the difference between all three clause types was found to be significant. Among the speakers in the Storyteller corpus there is only a significant two-way contrast between subordinate and all other clauses. In both cohorts of the combined Steelworker and Post-Industrial corpora, while the use of *be going to* has grown
in all three clause types, the contrast between subordinate clauses, main clauses, and apodosis has reached significance.

Figure 9.14, constructed identically to Figure 9.13 but with the predicted probability of *be going to* by year of birth and clause type, shows that through apparent time there has been a consistent division between the three contexts, and that they have been growing in parallel at a nearly constant rate. The parallelism may be masked at the extremities of the apparent time window as the number of data points decreases and the confidence intervals increases.

Grammatical person/animacy of the subject is a significant predictor for only the mixed Steelworker/Post-Industrial cohorts. In the Storyteller data non-animate subjects occur with *be going to* half as frequently as animate subjects, though this distinction could not be confirmed as significant. In the newer/younger mixed data non-animate subjects favour *be going to*, while first person subjects favour *will*. Other subjects are equally likely to occur with *be going to* and *will*. As with Figures 9.13 and 9.14, Figure 9.15 plots the predicted probability of *be going to* by year of birth, given both the fixed and random effects tested in Table 9.4. Here, though, tokens are divided by subject person and animacy. Figure shows that the probability of *be going to* has increased in first person and other animate subject contexts in parallel at a nearly constant rate across the full apparent time window. Among speakers born in the 1800s non-animate subjects are not distinct from other subject types, while among speakers born in the 1900s non-animate subjects are much more likely with *be going to* relative to other subject types. The probability of *be going to* in this context, once established, continues to grow at a constant rate in parallel with the two animate subject groups.
Predicted probability of be going to
by subject number and animacy and year of birth in Cape Breton English

Figure 9.15: Predicted probability of be going to by subject number and animacy and year of birth in Cape Breton English. Probability calculated based on a mixed effects logistic regression model of the full dataset with the fixed effects of sentence type, clause type, subject (grammatical person plus animacy), proximity, verb type, sex and centred year of birth, and the random effects of individual speaker and verb.

Proximity is only a significant factor among the older speakers in the mixed data, with proximate contexts favouring be going to and distal and unspecified contexts disfavouring. These two contexts, however, are not significantly different from each other. Figure 9.16 presents the predicted probability of be going to by year of birth, given both the fixed and random effects tested in Table 9.4, with proximate tokens coloured orange, distal tokens coloured green, and unspecified tokens coloured blue. Across apparent time distal tokens are more likely with be going to relative to proximate tokens. Unspecified tokens among speakers born after 1900 occupy the middle ground between these two contexts and all three rise in probability over relatively parallel trajectories at a constant rate through the apparent time window.

Finally, verb type is not a significant factor for any cohort.

9.7 Discussion

The patterns observed in the previous sections both confirm and challenge patterns presented in past literature — and perhaps account for some disagreement between studies.

The strong sentence type constraint aligns Cape Breton English with other Canadian varieties, and with British varieties in which be going to is categorized as "new and vigorous" — for whom negative and affirmative statements do not contrast significantly. If interrogatives were part of the initial pathway to grammaticalization (as Tagliamonte, Durham, and Smith 2014, 100–101 propose), the data from peripheral Cape Breton English adds to the substantial literature which shows that this initial pathway has had lasting profound effects on when be
predicted probability of be going to by proximity and year of birth in Cape Breton English

Figure 9.16: Predicted probability of be going to by proximity and year of birth in Cape Breton English. Probability calculated based on a mixed effects logistic regression model of the full dataset with the fixed effects of sentence type, clause type, subject (grammatical person plus animacy), proximity, verb type, sex and centred year of birth, and the random effects of individual speaker and verb.

be going to and will are selected in the grammar. As Figure 9.13 shows, even among speakers born prior to 1900 be going to is near categorical in interrogative contexts. Figure 9.13 also shows that negative and affirmative declarative contexts have consistently lagged behind this context, but have grown in parallel through apparent time. This contrasts with Tagliamonte, Durham, and Smith’s (2014) findings for older Scottish/Irish speakers for whom interrogative and negative declarative contexts patterned together, but does align Cape Breton with every other study of future temporal reference on a Canadian variety. Furthermore, Figure 9.13 shows the exact same parallel constant rate of growth in be going to probability for each sentence type as Denis and Tagliamonte’s (2017, 11) analysis of Toronto English instead of the complete reorganization of the sentence type found in Scottish/Irish English by Tagliamonte, Durham, and Smith’s (2014). Sentence type — the constraint with the greatest magnitude of effect — clearly points to a genetic similarity between Cape Breton English and other Canadian English varieties instead of between Cape Breton English and Scottish/Irish English.

While in each age/corpus cohort in the Cape Breton data subordinate clauses have the highest rate of be going to and apodoses have the lowest rate, it is only among the speakers from the combined Steelworker/Post-Industrial corpora that the difference between apodosis and other main clauses also reaches the level of statistical significance. This aligns the combined dataset with the unquestionably Loyalist-origin African–Nova Scotian speakers from North Preston and Guysborough Enclave reported in Poplack and Tagliamonte (2000) as well as speakers from Québec City, Montréal and Toronto (Torres Cacoullos and Walker 2009; Tagliamonte and D’Arcy 2009; Denis and Tagliamonte 2017).

Denis and Tagliamonte (2017), comparing their Toronto data with the data from the British Isles (Taglia-
monte, Durham, and Smith 2014), conclude that will is specializing for apodoses in varieties in which vigorous variation between forms is occurring; in older, more peripheral varieties, this specialization has not yet developed. In the Cape Breton data, unlike Toronto, this specialization of will for apodosis is not apparent. The slope of the binomial curve fitted to the apodosis data in Figure 9.14 has a positive slope, indicating increased probability of be going to over apparent time; however, in Denis and Tagliamonte’s (2017, Figure 5) Toronto data apodosis has a negative slope in a similarly constructed figure, indicating decreased probability of be going to over apparent time.

If the specialization of will for apodosis is a constraint that develops, rather than being present from the actuation of be going to for future temporal reference, it is possible that this development has not yet occurred in Cape Breton English. In this sense, Cape Breton English is similar, though unlikely genetically linked, to the peripheral Scottish and Irish communities studied by Tagliamonte, Durham, and Smith (2014).

With respect to proximity effects, the significant favouring of be going to in proximate context by the older Steelworker/Post-Industrial speakers aligns this data with Guysborough Village (Poplack and Tagliamonte 2000) and Québec (Torres Cacoullos and Walker 2009). Torres Cacoullos and Walker (2009) argue that the effect of proximity that they find in their data stems from an interaction between proximity, subject type and adverbial type in their data. Here, however, these factors have been controlled — with the highly divergent never and metadiscursive tokens removed from the analysis. Unlike Tagliamonte, Durham, and Smith’s (2014) findings in their combined corpus of British, Irish and Scottish varieties, for the data in this partition unspecified proximity aligns with explicitly distal rather than explicitly proximate contexts. Following the authors’ reasoning this would mean that the strong association between be going to and proximate contexts is real, and not simply an avoidance of be going to in distal contexts.

Across the three data cohorts in Table 9.4, explicitly proximate contexts occur more frequently with be going to than do explicitly distal contexts — though this effect only reaches the level of statistical significance among the older mixed cohort. In this regard the data from Cape Breton is in direct opposition to the data from Toronto examined by Denis and Tagliamonte (2017), where across apparent time explicitly distal contexts favour be going to. This seems unlikely given that elsewhere within the variable system (and for the other two morphosyntactic variables) Cape Breton and Toronto are in lockstep. The difference, however, may be methodological. In the Cape Breton data strongly will-favouring metadiscursive and never negation constructions have been excluded from the analysis. These contexts were not excluded in the Toronto data. As discussed in Sections 9.6.4.1, 9.6.4.2, and 9.6.4.5, all never tokens must be coded as non-proximate, while all metadiscursive tokens must be coded as proximate. Furthermore, in the present analysis future constructions without overt reference to when the future action will take place are coded as “Unspecified.” In Denis and Tagliamonte (2017) they are coded as distal and the binary contrast between distal and proximate (i.e., overtly specified to occur within 24 hours) is tested.

Consider Figures 9.17 and 9.18. In both figures the frequency of be going to and will in each data cohort is presented, but with data categorized as either proximate or non-proximate — future time overtly specified as taking place within 24 hours versus all other tokens. What is readily apparent is that completely different proximity effects manifest depending on how the data is manipulated. When the highly will-favouring metadiscursive and never-negation contexts are included, and overtly distal and unspecified contexts are grouped, an increasing favouring of be going to in non-proximate contexts emerges. Among the oldest speakers there is not a great difference between proximal and non-proximal contexts (29 percent vs. 26 percent be going to).23 By the youngest cohort the rate of be going to in non-proximal contexts increases to 49 percent, but in proximal contexts decreases to 22 percent. This pattern suggests increasing functional partitioning between will.

23. If tokens of never-negation are excluded the rate of be going to in not proximate contexts for the Storyteller corpus is 27 percent.
and be going to. It also aligns, at least in terms of the favouring of be going for non-proximal contexts, with Denis and Tagliamonte’s (2017) finding for Toronto. If the highly-will-favouring contexts are removed, but all non-proximal contexts are grouped there is yet another pattern: an early favouring of be going to in proximate contexts, but a levelling of this effect among the youngest speakers.

What 9.17 and 9.18 show is that metadiscursive contexts can have an extreme effect on the proximity effects observed in future temporal reference data. No other past study controls for this factor. The relative amount of metadiscursive constructions like I tell you... in each of the past studies datasets is unknown. Some datasets may have a relatively large number of metadiscursive contexts (as the Cape Breton data does), and for this reason be going to may appear to be disfavoured in proximate contexts (as in Figure 9.17). Other datasets may have a very small number of metadiscursive contexts, and for these datasets be going to may be more likely in proximate contexts (as in Table 9.4 and Figure 9.18). Without knowing the proportion of metadiscursive tokens in the Toronto, Québec City and Montréal, or British datasets it is difficult to compare Cape Breton English to varieties with respect to proximity effects.

![Expression of future temporal reference across corpora of Cape Breton English](image_url)

Figure 9.17: Distribution of expressions of future temporal reference by proximity in Cape Breton English — metadiscursive and never negation tokens included.

The lack of a verb-type effect is consistent with other studies on future temporal reference, including Poplack and Tagliamonte’s (2000) analysis of Ottawa English, Torres Cacoullos and Walker’s (2009) analysis of Québec and Montréal English, and Denis and Tagliamonte’s (2017) analysis of Toronto English. The finding is somewhat at odds with the significant favouring of verbs of motion found by Poplack and Tagliamonte (2000) for the two African–Nova Scotian communities and Guysborough Village. Perhaps Industrial Cape Breton, even older speakers from Industrial Cape Breton, align more closely with extra-provincial varieties than do these smaller Nova Scotia communities.

In summary, with respect to overall rate of be going to use, Cape Breton English more closely resembles other transported varieties rather than British, Irish or Scottish English. Across the full Cape Breton dataset there is a significant two-way contrast between questions and statements like other Canadian varieties, but there is not
Figure 9.18: Distribution of expressions of future temporal reference by proximity in Cape Breton English — metadiscursive and never negation tokens excluded.

a significant distinction between negative and affirmative statements, like conservative British/Irish/Scottish varieties. With respect to every other tested constraint on variation between Cape Breton and Toronto, the two are identical. The two precisions to this statement, however, are that while apodosis disfavours be going to relative to other clause types in both communities, it is only in Toronto that this contrast is increasing over apparent time, and that Cape Breton matches Toronto’s proximity effect only when the data is organized using a, I suggest, suboptimal coding scheme.

9.8 Conclusion

The expression of future temporal reference in Cape Breton English adds to the literature on changes to the English future and may prove useful for future studies for several reasons. As Tagliamonte, Durham, and Smith (2014) point out, while be going to has been attested in the literature as early as the 15th century, these attestations are mainstream varieties and drawn from written sources. The preceding analysis is the largest study to date using spoken vernacular data from a single community, and has the largest time depth in real/apparent time of any such study. The analysis has shown a consistent, albeit slow, growth of be going to across the 20th century in line with its long development within the language. Cape Breton’s status as a peripheral variety within a Canadian context suggests that the community ought to exemplify an older version of the language — much as the peripheral communities studied by Tagliamonte (2013a) and Tagliamonte, Durham, and Smith (2014) do. However, with respect to future temporal reference (as with stative possessives and deontic modality verbs) Cape Breton has progressed in lockstep with Toronto in terms of rates and grammatical patterns of use. Thus rather than a window into the past of be going to, the current analysis should be considered as a carefully considered
analysis of be going to and will’s present competition within English.

What the current analysis has shown is the strong and persistent effect of subject and clause type on the choice between the two forms. Be going to is believed to have grammaticalized via questions and subordinate clauses, and these two specific contexts reflect that entry point by exerting a strong conditioning force on the selection of be going to over will. Through real and apparent time additional sentence types (negative and affirmative statements) and other clause types (main clauses, including apodosis) show an increase as well, suggesting a consistent parallel growth of be going to across contexts. That said, there are some changes that do occur to the system. For example, the use of be going to and will with different subject types appears to change among speakers born early in the 20th century.

This analysis has also provided yet another piece of strong evidence, along with Chapter 8, that perlocutionary discourse markers behave atypically within the grammar and are sites of relic conservation. The careful pruning of data prior to regression analysis in order to set these types of tokens aside is crucial because they are both frequent and can exert a substantial influence on the interpretation of variable patterns. Expounding on the nature of how perlocutionary markers fit into a theory of grammar is beyond the scope of this paper; however, Jackendoff (1996, Ch. 7) proposes that fixed expressions, clichés, idioms, etc., might be stored as “chunks” akin to compounds within the lexicon while the individual components of the fixed expressions, clichés, idioms, etc. may also be stored in the lexicon and used compositionally elsewhere. This is not unlike the initial stage of grammaticalization of previously compositional phrases like be going to or be like, etc. If constructions like I’ll tell you, or I must admit, when occupying the left-edge as a perlocutionary marker, become conventionalized as fixed expressions and are stored in the lexicon, innovations like be going to or have to may take longer to penetrate them. For this reason, in future studies of linguistic innovations, it may be wise to examine patterns of change in different discursive contexts independently, or at the very least operationalize discursive context as a potential constraining factor.
Part IV

Discussion and Conclusion
Chapter 10

Overall Discussion

10.1 Overall Summary

In the first chapter of this dissertation the speech community of Industrial Cape Breton was introduced, so too were the numerous examples from the linguistic literature and popular fiction that have labeled Cape Breton English as a specifically Scottish-derived or Gaelic-influenced variety. This chapter further outlined the four sociolinguistic variables to be used as test cases for assessing the genetic link between Cape Breton English and Scottish English. The method of comparison was comparative sociolinguistics, which looks not just at the frequency of variants but also the grammatical rules underlying variation. This method relies on the distinction between linguistic transmission and linguistic diffusion (Labov 2007), the former of which is characterized by the replication of grammatical patterns through childhood acquisition, the latter of which is characterized by the simplification of grammatical patterns as a consequence of adult learning. *Be like* was selected as a prototypical diffused variant and was used to show the kind of simplification of grammatical patterning that occur via diffusion. The other three variables — stative possession, deontic modality, and future temporal reference — were selected because these three variables have very different grammatical patterning in North American and Old World Englishes.

Chapter one also introduced the sociolinguistic issues relevant to this research, which include the role that social factors like age, sex, and social status and geographic factors like centrality and peripherality play in variation. This chapter further introduced the broader sociolinguistic questions of whether innovations in the grammar result in layering or functional partitioning of variants.

Chapter two outlined the history of the planting of English on Cape Breton Island and the subsequent language shift of later Gaelic-speaking settlers. It further discussed how industrialization and deindustrialization led to the coalescence of a Cape Breton identity, and how the evaluation of non-standard speech as being inherently local to Cape Breton relates to that identity. A thorough cataloguing of previous research on Cape Breton English and its divergent forms revealed numerous lexical and phonetic features that set Cape Breton apart from Inland Canadian speech; however, few divergent morphosyntactic features could be identified (with the moribund *after*-perfect perhaps being the best exception).

Following these introductory chapters the four cases studies were presented. The first, an analysis of the adoption of the *be like* quotative in Cape Breton revealed, as predicted, clear signs of diffusion: a diachronic time lag and a simplified subject and gender constraint in Cape Breton relative to other varieties. Furthermore, this chapter provided corroborating evidence for Tagliamonte and Denis’s (2014) proposal that in some communities
the *be like* quotative can be diffused lexically, with subsequent generations inferring grammatical patterns. It also suggested that *be like* is diffusing in Canada following a wave or gravity model.\footnote{The wave and gravity models of geographic diffusion of linguistic change make different predictions; however, in both scenarios, assuming eastward diffusion, *be like* would be predicted to diffuse to Quebec City sometime after it had diffused to Toronto, and to Cape Breton and to St. John’s sometime after it had diffused to Quebec City.}

The second case study explored variation in the stative possessive system. The major finding was that Cape Breton and Ontario possess nearly identical variable patterns. This led to the conclusion that the two varieties share a genetic link, and that constraints governing the choice between *have, have got,* and *got,* must pre-date the divergence of inland and coastal Loyalist speech. The following chapter, which examined the deontic modality system, showed not only this same evidence for genetic link, but also showed that both the stative possessive and deontic modality system in Cape Breton share a common subject type constraint — a subject type constraint only attested in other Loyalist-founded speech communities (Toronto and southern Ontario). This was considered very strong evidence for the transmission of grammatical patterns, rather than later geographic diffusion.

The final case study, which tracked the rise of *be going to* in Cape Breton English, again showed rates of *be going to* and variable rules for its use that align with Inland Canada and not with the British Isles. This chapter presented a number of methodological precisions (namely the careful consideration of discourse context) for the study of future temporal reference.

In these four chapters is a discussion of the major findings of each chapter’s case study. In the remainder of this chapter I will discuss four issues that arise when these four studies are considered collectively.

### 10.2 Being local and global

In Section 8.7.1 I discussed how the diachronic suppression of *have got* and *have got to* in Cape Breton English in favour of possessive and deontic modality systems that only include *have/have to* and *got/got to* provided strong evidence for D’Arcy (2015) proposal that *have* has become fully lexical in Canadian English. The changes to the possessive and deontic modality systems, however, are reinforced by the social embedding of these variants. The reorganization of both systems — whereby the predominant two options become an externally prescribed change from above and a stigmatized form — reflect, for example, the dual pressures of Cape Breton English speakers to converge with mainstream norms but also to express local affiliation (as discussed in Section 3.1.2).

While speakers readily align with external norms, in contexts where they choose *not* to employ the prescribed variant *have,* the variant they increasingly employ is stigmatized *got* rather than the more neutral *have got.*

Consider how men and women in Cape Breton have used the variants of these two variables across real and apparent time.

For both variables women consistently use very high rates of prescriptive *have/have to* in line with Labov’s (2001, 266, and elsewhere) “gender paradox” whereby women hew closer to linguistic norms when those norms are overtly prescribed. If Cape Breton English is rapidly converging with mainstream Canadian norms, the speakers who are doing that converging are male speakers. From the 1880s to the 1940s Cape Breton men quickly go from being majority *have got/have got to* to matching the rates of *have/have to* use of Inland Canada. Among both men and women born after 1950 there is relative stability as the change towards *have/have to* slows down.

But what do these speakers do when they choose *not* to adhere to prescriptive norms? They choose the stigmatized *got/got to* forms instead of the *have got/have got to* forms. This is not an example of retrograde change, because the overall rate of secondary variants remains stable, and older speakers in the community do not employ the *got/got to* forms with much frequency — at least not relative to either *have got/have got*
Figure 10.1: Distribution of stative possession verbs by sex in Cape Breton English — male tokens grouped by decade of birth, female tokens in 20-year groups due to smaller number of tokens.

Figure 10.2: Distribution of deontic modality verbs by sex in Cape Breton English — male tokens grouped by decade of birth, female tokens in 20-year groups due to smaller number of tokens.
to of have/have to. The elevation of the stigmatized tertiary variant occurs exactly because it is stigmatized. As discussed in Chapter 3.2, non-standard language features are confounded with local language features in Cape Breton. In a linguistic milieu where rapid convergence with external norms has occurred, the use of stigmatized/local vernacular forms can act to signal affiliation and may garner covert prestige. The adoption of specifically stigmatized variants, or the use of variants in only stigmatized linguistic contexts is a phenomenon observed in other peripheral North American varieties that are increasingly accommodating to external norms (e.g., Childs et al. 2010; Childs and Van Herk 2014; Schilling-Estes 2003)

It is important to note that Cape Breton male speakers are not whole-heartedly rejecting off-island mores — their use of stigmatized forms is only in contexts where the majority form is not employed, and the majority form is employed at rates that are comparable to Toronto, etc. In fact, with respect to all four changes in progress men are either the current leaders of change towards off-island norms (be going to) or indistinguishable from females with respect to the adoption of those norms (be like, have, have to). That men show the most dynamic language change and overall the closest hewing to external norms when the vast majority of the male speakers in the data are steelworkers with no or little post-secondary education is astounding. In fact, I do not have a satisfactory answer to explain the phenomenon other than to suggest that within a Cape Breton context women are considered to be torchbearers of local custom and identity — an idea explored by Gardiner-Barber (1996; 2002) — and thus overall more likely (at least in the second half of the 20th century) to be sociolinguistically conservative.\footnote{Though, Cheshire, Kerswill, and Williams (2005) argue that Labov’s (2001) three principles of linguistic change vis-à-vis gender have only been adequately observed for specifically sound change. “There are no such general principles for morphosyntactic changes, and we know still less about the social embedding of changes at higher levels of structure. The few reports that do exist give a contradictory picture. We [argue] that the relative infrequency of syntactic variants makes them unlikely to occur with sufficient frequency to become habitually associated with the speech of either women or men. This in turn mean that there is no reason to suppose that syntactic features will follow similar patterns of change to phonetic and phonological variables,” (Cheshire, Kerswill, and Williams 2005, 144). That being said, got, gotta, and gonna are clearly prescriptively stigmatized in English, and stigmatized/informal or covertly prestigious speech may more generally be associated with males (e.g., Trudgill 1972).}

The other explanation is that the observed male sociolinguistic dynamism is a reflection of the economic and

---

Figure 10.3: Distribution of be like by sex in Cape Breton English — male tokens grouped by decade of birth, female tokens in 20-year groups due to smaller number of tokens.
Figure 10.4: Distribution of *have* by sex in Cape Breton English — male tokens grouped by decade of birth, female tokens in 20-year groups due to smaller number of tokens.

Figure 10.5: Distribution of *have to* by sex in Cape Breton English — male tokens grouped by decade of birth, female tokens in 20-year groups due to smaller number of tokens.
demographic upheaval that occurred in Cape Breton in the latter half of the 20th century with the decline of heavy industry, described in Section 3.1.2. The social disruption this decline caused may have affected men more profoundly than women — men, for example, were much more likely to lose their jobs as steelworkers or miners and were forced to re-train for other industries or seek industrial or primary resource work in other locales. This may have resulted in Cape Breton men being particularly sensitive to sociolinguistic variation across industries and they have felt additional pressure to moderate their use of conservative/local/non-standard and embrace off-island ways of speaking. This harkens to Milroy’s (1980) study of Belfast, where in neighbourhoods like the Hammar and Clonard, out-of-work men were required to look for work outside of the community and perhaps because of this employed a higher rate of standard forms than females.

Taken together, these two explanations may explain the relative similarity in rates of use of all four innovative variants. Men in Cape Breton are slightly more linguistically innovative/standard than might be expected, while females are slightly more linguistically conservative/non-standard than might be expected, resulting in convergence between sexes. This does not mean that the social categories of “male” and “female” are not important within Cape Breton — simply that they are not reflected in the variation between the four variables studied here.

10.3 Cape Breton English as Canadian English

While the field has generally come to agree about the Loyalist origin of Canadian English, there was been some historic debate among linguists about the relative importance of different dialect groups to the creation and nature of Canadian English. There were two main positions: Canadian English is most importantly shaped by Loyalists, with dialect differences between Inland Canada and the Maritime reflecting dialect differences between inland and coastal Loyalists who evacuated the United States via different routes; and Canadian English is most importantly shaped by the massive amount of (especially) 19th-century immigration to the country by speakers from the British Isles, most notably Northern Britain, Scotland, and Northern Ireland. The former is perhaps
best exemplified by M. Bloomfield (1948, 6) (though see also Avis 1973; Chambers 1998; Dollinger 2008), who writes that “Canadian English, as has been recognized by some observers, is to all intents and purposes General American with a few modified sounds, usually parallel in American sub-dialects, and with some vocabulary variation.” The latter, perhaps best exemplified by Scargill (1957), who claims the “Loyalist theory” is based on potentially misguided assumptions (Scargill 1957, 13). Scargill’s (1957) position, (Dollinger 2008, 128) calls it “Numerical Swamping”,3 argues that the sheer number of post-Loyalist 19th-century immigrants from the British Isles “must be given some consideration.”4

Cape Breton English is perhaps one of best test sites for disputing Scargill’s (1957) “numerical swamping” claim. Cape Breton had a very small Loyalist base that was swamped by an overwhelming Scottish majority. While it is true that, relative to Standard Canadian English, there are divergent features in the variety (originating from multiple input sources, including New England, Newfoundland, mainland Nova Scotia, Ireland and Scotland), Cape Breton English has maintained stative possession and deontic modality systems that are identical to other Loyalist-input Canadian varieties. That the oldest, most rural speakers in the Cape Breton data use the most have got/have got to forms — unquestionably Yankee, and only just diffusing to the Northern British Isles — shows that at least two of the linguistic norms of one Loyalist-founded dialect area did not succumb to the influence of the later, greater 19th-century immigration from the British Isles.

If the propagation of Loyalist patterns in Cape Breton, despite numerous divergent features, can be generalized to other rural communities in Ontario and the West, Cape Breton can serve as a model for how Canadian English homogenized (i.e., coalesced to Loyalist patterns) in those rural communities.

### 10.4 Layering versus functional partitioning

Section 2.4 introduced two possible results of innovation in the grammar: layering and functional partitioning. As described in Section 2.4 layering is part of the process of grammaticalization, whereby nonce or formerly lexical expressions take on grammatical meanings that overlap with already existing functional forms (see, for instance Hopper 1991, 22; Hopper and Traugott 2003, 125). Grammaticalizing forms may begin in restricted contexts, but gradually expand to all contexts where the forms over which they have been layered are used. Over time the innovative form may increase in frequency of use and the older form may become obsolete. Functional partitioning occurs when a grammatical variant is introduced into the system and the new form and the older form become specialized for specific syntactic, semantic, or pragmatic functions (Hopper and Traugott 2003, 125).

Across the four variables investigated evidence of both layering and functional partitioning is present. For the quotative system, be like appears to be saturating the linguistic system, suggesting layering. The few times where be like is not used by the youngest speakers are contexts where more semantically rich variants (like holler, yell, or whisper) are required. The former semantically neutral variant say appears to be becoming specialized semantically for describing how a speaker pronounces something, or whether or not they use a particular expression, as in (33) on p. 68. This aligns with Wallenberg’s (2013) view that specialization only occurs when speakers attribute contrastive meanings (denotational or otherwise) to competing forms.

Both the future temporal reference and deontic modality systems show layering with putative specialization of deontic must for metadiscursive perlocutionary markers — though this terminal refugium is being invaded diachronically by majority variant have to. While not exactly the same as specialization based on semantic

---

3. After Lass (1990)
10.5 In search of Gaelic influence

One question remains: why has not Gaelic exerted more of a lasting influence on Cape Breton English?

Given the huge influx of Gaelic-speaking Scots to Cape Breton, as well as Sydney’s status as an early 20th century locus of ethnic diversity, the expectation might be that Cape Breton English would have many more divergent speech features than those actually recorded. The imagined version of Cape Breton English held in the minds of both local and non-local residents is certainly divergent.

I suggest that the reason Gaelic has not exerted a stronger substrate effect on Cape Breton English is simply that it was for many years unacceptable for public use, especially among schoolchildren, as discussed in Sections 3.1.1 and 3.1.2.

The consequence of industrialization in Cape Breton coincides with Schneider’s (2007) third stage of post-colonial English development: nativization, which Schneider (2007, 40) describes as “the central phase of both cultural and linguistic transformation.” This consequence was the development of a specifically-Cape Breton identity formed in opposition to off-island industrial and political authorities and via the shared hardship of “making do” (Gardiner-Barber 2002, 400-401). Like other colonial Englishes the process of identity re-writing in an establishing community is fundamental to understanding how the English variety in that community manifests. By the time of industrialization and its associated identity re-writing the Gaelic-language in Cape Breton had endured over 75 years of institutional suppression and social disadvantage. The “local” variety of English that develops during this period — that is, the version of English that is viewed as specific to Cape Breton by residents — is therefore not one that is highly influenced by Gaelic transfer, nor one that conforms to external standards, but instead one that is built from the working-class Loyalist, Newfoundlander, Irish, etc. vernacular Engishes spoken in the coal pits or blast furnace, over tea in company houses, along picket lines, and especially in the playgrounds. The face-to-face interactions between speakers aiming to build and reinforce local social ties work over generations to homogenize the local vernacular (see Chambers 2002). Some sounds, some words, some discourse rituals, and perhaps some morphosyntactic features may be incorporated into this vernacular variety, but the bedrock is still the founding version of English. If these extra features are evaluated as saliently non-standard, they may be eagerly adopted by a community that envisions itself as existing in opposition to off-island hegemonic social, political, and economic institutions. Cape Breton English is therefore shaped not just

---

5. This is less so the case with Acadian French and Mi’kmaq in the areas of the island in which speakers of these languages are dominant, e.g., Chéticamp, Arichat, and the five Mi’kmaq First Nations communities.

by a founding costal Loyalist English input, but by a bevy of vernacular Englishes — L2 Gaelic or otherwise. The legacy of this is the inclusion of Newfoundland, New England, Scottish and Irish-origin vernacular lexical items in the *Dictionary of Cape Breton English* (Davey and MacKinnon 2016), *Da Mudder Tongue* (Gray 2006; 2007), and on Cape Breton souvenirs like Photograph 2. The perpetuation of these vernacular forms being iconic of Cape Breton English was also reinforced and perhaps codified — consistent with Schneider’s (2007) third step, endonormative stabilization — during the Cape Breton cultural renaissance of the late 1960s to early 1980s when Baby Boomer performers took their working-class parents as models for prototypical local speech (e.g., Westhaver 1996, 93-94). This included salient lexical, phonetic/phonological, morphosyntactic and discourse-pragmatic features. “Authentic” Cape Breton speech is thus imagined as Celtic, founded on Loyalist English, influenced by other vernacular varieties, and among speakers born after WWII, something you “put on” or perform. This is perhaps why speakers in the *Post-Industrial* corpus born in the 1980s — the group of speakers who chose to stay in Cape Breton when so many of their peers have left (Statistics Canada 2012) — employ salient vernacular forms (e.g., *got*, *got to*) when performing local speech, as in (52) on p. 94, and perhaps to index a localness more generally.
Chapter 11

Overall Conclusion

An important lesson to be learned from the analyses presented by this dissertation’s case studies is that the “rapid convergence” of Maritime Canadian English to mainstream Canadian norms (Keifte and Kay-Raining Bird 2010, 62) — here represented by ostensibly the most divergent Maritime regional dialect — is in fact one of quantity not kind, with the aligning of the frequency of variants’ use rather than grammatical constraints.

While the option to use certain phonetic/phonological features\(^1\) and lexical items\(^2\) may continue to differentiate Canadian Maritime English from mainstream Canadian English, with respect to deep-rooted grammatical patterns like the variable rules for using have, have to and be going to, the two varieties are indistinguishable. Cape Breton English, thought by some to be an exception to Canada’s English-language homogeneity, instead provides strong evidence for homogeneity resulting from a persistent Loyalist founder effect. The similarity between Inland Canadian English and Cape Breton English with respect to these variables is due to the communities’ common initial English input, rather than eastward diffusion.

That being said, it is not as if diffusion does not occur. The analysis of be like shows strong evidence that the feature was first used both later and with less complex grammatical conditioning in Cape Breton compared to Toronto. These facts point to the likely diffusion of be like to Cape Breton English from further west. While some communities worldwide show evidence of parallel instantaneous development of be like, as perhaps a reaction to an epiphenomenal change within the grammar (as argued by Gardner et al. 2016), or because of heretofore undocumented rapidity in diffusion spurred by the mid-twentieth-century explosion of global youth culture and connectivity (according to Tagliamonte, D’Arcy, and Rodríguez-Louro 2016), Cape Breton is not one of those communities. The first robust users of be like are speakers born in the 1980s. All older be like users either work closely with teenagers or have lived off-island in be like-using communities. Be like did not take root in Cape Breton among the Baby Boomers, but instead among their children — a generation later than elsewhere in Ontario and Québec.

The status of Québec as a linguistic barrier between Atlantic Canada and Inland Canada has been called into question. Dion and Poplack (2005; 2007) have argued that be like arose in Québec English via drift or extralinguistic means (e.g., the media). While the authors are not alone in suggesting that be like can arise through drift or spread via lexical diffusion (possible via extralinguistic avenues), given the timing and grammatical patterning of be like in Québec, Cape Breton, and St. John’s, wave-like geographic diffusion is also quite likely. The striking similarity between Toronto, Ottawa, Québec City/Montréal, African Nova-Scotian communities, and Cape Breton with respect to future temporal reference also points to a cohesive and unbroken dialect region.

---

1. E.g., pre-rhotic low vowels, coat–coach merger, slit fricatives, ingressive yes and no, etc.
2. E.g., sneakers, sook, etc.
from a morphosyntactic perspective.

Based on historic and — presented here for the first time — linguistic evidence, Cape Breton English owes much of its character to its Loyalist input, perhaps more so than the influence of post-Loyalist Newfoundland, Irish, or Scottish settlers. The latter’s influence is not insignificant — there are several phonetic/phonological features, at least one morphosyntactic and perhaps a handful of lexical features inherited from Gaelic and/or shaped by Gaelic/Scottish cultural influence. The presence of these features should not, however, be evidence that Cape Breton English is unrelated to mainstream Canadian English or “more or less identical with the English of the Scottish Highlands” (Trudgill 2004, 7). Using the comparative sociolinguistic method, coupled with a careful combing of 19th-century Nova Scotia literature and a historic account of Cape Breton’s settling, I argue strongly that Cape Breton English is not “more or less identical with the English of the Scottish Highlands” but instead, for three morphosyntactic variables, more or less identical to the rest of Canada. This similarity points to a genetic relationship between Cape Breton English and Inland Canadian English.

Recognizing this relationship has both linguistic and social implications. The belief that Cape Breton’s distinctive ways of speaking result from a preservation of Scottish forms and exists in opposition to hegemonic external norms helps to build and reinforce local identities, yet the linguistic classification of Cape Breton English as external to — rather than on the periphery of — mainstream Canadian English works to delegitimize the dialect and reinforce the notion that Cape Breton English speakers are not one-and-the-same as mainstream Canadian speakers, extrinsic to their shared history, linguistic or otherwise. If educators or policy makers look to linguists for guidance with respect to the status of vernacular Cape Breton forms and are presented with unsubstantiated claims like Cape Breton English represents “a failure of initially non-Canadian varieties…to assimilate fully to the Canadian English spoken in the larger regions around them” (Boberg 2010, 28) those vernacular forms become deficiencies for speakers. Their vernacular dialect heritage becomes something needing to be overcome. This is highly inequitable considering Cape Breton English and mainstream Canadian English share the same roots.

If educators and policy makers are informed that based on strong linguistic research, Cape Breton English is unquestionably a regional form of Canadian English — as Toronto, Vancouver, or Calgary’s dialects are — that originates from the same Loyalist wellspring as mainstream Canadian English, then Cape Breton vernacular forms become part of a larger negotiation of what constitutes standard Canadian English and Cape Breton’s linguistic features are legitimatized, as are their users.

The full Corpus of Cape Breton English offers perhaps one of the most intriguing views of the history of Canadian English. Among its oldest speakers, born in the 1870s through the 1910s, there is evidence of a very familiar yet very different version of Canadian English. The youngest speakers in the corpora, who were still in middle and high-school when they were recorded in 2009-2011, exemplify Canadian English at the beginning of the 21st century. These speakers employ new variants (be like) readily and for changes in progress with long time depths (e.g., be going to) employ increasingly frequent variants with the same rate as their big city counterparts. In this sense the Corpus of Cape Breton English is both a window into the past, and a snapshot of the present; a wooly mammoth and member of the vanguard of Canadian English.
Part V

References and Appendices
Bibliography


———. 1940. Linguistic Geography. Queen’s Quarterly 47:38–47.


BIBLIOGRAPHY


Durian, David. 2012. A New Perspective on Vowel Variation across the 19th and 20th Centuries in Columbus, OH. Ph.D. dissertation, The Ohio State University.


Gardner, Matt Hunt. 2009. Frontin’: Low back vowels and six Nova Scotians living in St. John’s. Presented at the 32nd Annual Meeting of the Atlantic Provinces Linguistic Association (APLA), Dalhousie University, Halifax, NS.

———. 2010a. Exodus: iconic speech features and identity amid Cape Breton Island’s rapid population decline. Presented at Sociolinguistics Symposium 18, University of Southampton, Southampton, UK.


———. 2013b. Variable word-final (t,d) in Cape Breton English. Generals Paper, University of Toronto.


Gardner, Matt Hunt, Derek Denis, Marisa Brook, and Sali A. Tagliamonte. 2013. The new global flow of linguistic influence: *be like* at the saturation point. Presented at New Ways of Analyzing Variation (NWAV) 42, University of Pittsburgh & Carnegie Mellon University, Pittsburgh, PA.

———. 2016. From the bottom to the top of the S-curve: ‘be like’ and the Constant Rate Effect. Presented at the 90th Annual Meeting of the Linguistic Society of America, Linguistic Society of America, Washington, DC.


Henley, Bridget, Meghan Hollett, Elizabeth Ingram, and Gerard Van Herk. 2008. Grey’s Anatomy viewers as a community of choice (seriously!) Presented at New Ways of Analyzing Variation (NWAV) 37, Rice University, Houston, Texas.


Macgillivray, Don, and Brian Douglas Tennyson, eds. 1980. Cape Breton Historical Essays. Sydney, NS: College of Cape Breton Press.


Murray, Lindley. 1817. *English grammar: Adapted to the different classes of learners. with appendix, containing rules and observations for assisting the more advanced students to write with perspicuity and accuracy*. York: Thomas Wilson & Sons.


Noble, Shawn. 1985. To *have* and *have got*. Washington, D.C.: Presented at New Ways of Analyzing Variation (NWAV) 14, Georgetown University, Washington, DC.


Quinn, Heidi. 2004. Possessive have and (have) got in New Zealand English. Presented at New Ways of Analyzing Variation (NWAV) 33, University of Michigan, Ann Arbor, MI.


Wallenberg, Joel C. 2013. A unified theory of stable variation, syntactic optionality, and syntactic change. Presented at the 15th Diachronic Generative Syntax Conference, University of Ottawa, Ottawa, ON.


Appendix A

Additional Figures and Tables

Figure A.1: Conditional inference recursive partitioning tree — have vs. have got/got in Cape Breton English. Parameters included: year of birth, sex, corpus.
Figure A.2: Conditional inference recursive partitioning tree — have to vs. must/have got to/got to in Cape Breton English. Parameters included: year of birth, sex, corpus.

*ST = Storyteller corpus, SW = Steelworker corpus, PI = Post-Industrial corpus.
Table A.1: Mixed effects logistic regression— contribution of factors to the probability of *have to* for expressing deontic modality in objective obligation contexts in Cape Breton English.

<table>
<thead>
<tr>
<th>Deontic modality <em>have to</em> in Cape Breton English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>aic</strong></td>
</tr>
<tr>
<td><strong>input</strong></td>
</tr>
<tr>
<td><strong>total N</strong></td>
</tr>
</tbody>
</table>

**Fixed Effects**

**Subject Type**

<table>
<thead>
<tr>
<th>Type</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun Phrase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd Person (plural)</td>
<td>66</td>
<td>82</td>
<td>222</td>
</tr>
<tr>
<td>1st Person</td>
<td>.43</td>
<td>72</td>
<td>436</td>
</tr>
<tr>
<td>2nd Person</td>
<td>.45</td>
<td>70</td>
<td>1041</td>
</tr>
<tr>
<td>3rd Person (singular)</td>
<td>.33</td>
<td>70</td>
<td>320</td>
</tr>
</tbody>
</table>

**range**

<table>
<thead>
<tr>
<th>Sex</th>
<th>FW</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>[.56]</td>
<td>81</td>
<td>169</td>
</tr>
<tr>
<td>Male</td>
<td>[.44]</td>
<td>72</td>
<td>2843</td>
</tr>
</tbody>
</table>

**Year of Birth**

*continuous factor*  
+0.034 log-odds

**Random Effects**

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.9</td>
<td>188</td>
</tr>
</tbody>
</table>

Sum contrast coding.

Aic, input, and factor weights for significant factor groups derived from most parsimonious model. Factor weights for non-significant factor groups derived from full model.

Factor weights converted from log-odds.

\[ p = \frac{e^x}{1+e^x} \]

Factor weights with non-significant contrasts indicated with [  ].

Correlation of fixed effects: \( r < |0.063| \)

Additional treatment contrast coding used to test orthogonal contrasts for parameters (factor groups) with more than two levels (factors) using most parsimonious model.

Significant contrasts indicated with horizontal lines.
Appendix B

Bungi

There exists perhaps some evidence of historical stative possessive got forms in a form of Scottish-Canadian English. In Bungi, a 19th-century Red River Valley, Manitoba, dialect (Bakker and Grant 1996, 1115; Stobie 2010, 207), or post-creole (Pentland 1985), of Canadian/Scottish English that was lexically and grammatically influenced by Ojibwe, Cree, Salteaux, Lowland Scots, Orcadian, Shetlandic/Norn, and especially Scottish Gaelic, there are attestations of got forms with a stative possessive meaning. In Blaine’s (1989) description of Bungi there are several transcribed utterances from a heritage Bungi speaker born in the 1910s with a stative possessive be got form. The author, in discussing be got, gives the following examples (189):

(92) a. "They’re got new railing on the bridge.” (observed)
b. “That’s a new fence they’re got.” (observed)
c. "You’re got your tea, my girl.” (observed)
d. “We’re got lots of time, my girl.” (observed)
e. “I’m not got that big money.” (observed)
f. "I’m not got my glasses.” (observed)
g. “Auntie, I’m got the sinus.” (recorded)
h. "Comforters and everything, now, we’re got.” (recorded)
i. "You’re got brains and wits you know.” (recorded)

The same speaker also uses have and have got, displaying the same stative possessive variants one would expect to find among any older Manitoba speaker recorded in the 1980s (based on Tagliamonte and Denis 2014).

(93) a. “I have my sister, Susan, that’s seventy-five.” (Blaine 1989, 151)
b. “I have a creamy colour [pantyhose] home.” (Blaine 1989, 170)
c. “...same as they have over to Maggie Town.” (Blaine 1989, 173)
d. "Under I’ve got the woollen shirt and the bra.” (Blaine 1989, 148)

Bungi shares many lexical, grammatical and even discourse features with Cape Breton English (c.f., Falk 1989) — as well as Newfoundland English, AAVE, and other English varieties; however, there is no be got in any of the Cape Breton English data. Blaine (1989, 189–191) traces the roots of this feature to Orcadian using historical text data. While Orcadian is spoken in Scotland, it derives from Scots — a language/variety not present in the Scottish Highlands or Hebrides (e.g., Wells 1982b, 395), from where Cape Breton’s Scottish population originated. So, despite both being Canadian, and both having Scottish/Gaelic influence, Bungi and Cape Breton’s stative possessive systems are likely unrelated.