TRANSMISSION AND DIFFUSION ABOVE THE LEVEL OF PHONOLOGY: EVIDENCE FROM THUNDER BAY

by

Derek Denis

Forum paper submitted in conformity with the requirements for the degree of Master of Arts
Graduate Department of Linguistics
University of Toronto

© Copyright by Derek Denis 2009
Abstract

This paper examines Thunder Bay English (Roeder 2008), by focusing on two variable systems known to be changing in Toronto: the deontic modality system, as in (1) (Tagliamonte & D’Arcy 2007a); and, the quotative system, as exemplified in the (2) (Tagliamonte & D’Arcy 2007b).

(1) I have to be home by nine. I gotta get to bed.
(2) When I got it, I’m like, “Oh that’s what they were talking about

These variables are used to explore Labov’s (2007) distinction between transmission and diffusion as explanations of linguistic change. While the former is characterized by the “faithful transmission” of the underlying grammar constraining variable systems (Labov 2007:346), Labov (2007:347) observes the diffusion of linguistic innovation involves the imperfect replication of linguistic rules and constraints. Variationist methods can apply this distinction in order to investigate the course of linguistic change.

Results of a quantitative analysis of Thunder Bay English reveal that the underlying grammar of the deontic modality system is constrained in the same way observed in Toronto suggesting transmission; the change likely arrived in Thunder Bay with its settlement. However, quotative be like must have been diffused, as it is absent in the over 30 population. Demographic evidence strongly suggests that Toronto is the source of be like to Thunder Bay but some constraints observed to be active are not consistent with the Toronto system. This follows from Labov (2007) but leaves unanswered the question of why certain constraints and not others are successfully acquired through diffusion

While considering the methodological consequences of Labov (2007) (e.g. is the transmission/diffusion distinction consistent above the level of phonology?), the perspective from Thunder Bay adds to the growing body of variationist findings on Canadian English.
Acknowledgements

There are a lot of people to thank – family, friends, animals and linguists alike.

Linguists: My fellow graduate students at the University of Toronto – past, present and future. I am honoured to place this paper in the first bookcase of the back room amongst your works and especially to my current classmates with whom I share the department’s thickest binder yet. Thank you to Alex Motut, Alice Meyers, Ann Dolan, Anne Johnson, Brenna Haimes Kusumoto, Jackie Peters, Jen Higgs, Joanna Chociej, Kaori Furukawa, LeAnn Brown, Lizzie DiGiacomo, Mercedeh Mohaghegh, and Suzi Dumitrescu-Georgescu. I will truly miss sharing ideas, bagels, and coffee (and drinks) during Forum. There are a number of other students who I need to give written fist bumps to - Dylan Uscher, Ryan McIsaac, Cori Hanson, Isaac Gould, Nik Rolle, Cathleen Waters, Bridget Jankowski, Sandrine Tailleur, Michelle St-Amour, Ailis Cournané, Kenji Oda and Richard Compton. I need to thank Kevin Heffernen and Becky Roeder for a lively discussion about how to properly get through grad school over beers in San Francisco. Thanks to Keren Rice, Susana Béjar, Naomi Nagy, Yoonjung Kang, and Elan Dresher for letting me learn from them this year and to Elaine Gold, Alana Johns, Elizabeth Cowper, Chiara Frigeni, Cristina Cuevo, Alexei Kochetov, Ron Smyth, Becky Roeder, Heriberto Avelino and Mohammad Vehdi for letting me learn from them as an undergrad. Thank you also to Andrew Nevins, Salikoko Mufwene, David Adger, and Peter Siemund who visited the University this year and shared their work and ideas with me. Thanks to Gerard van Herk, Stefan Dollinger, James Walker, Shana Poplack, Nathalie Dion and everyone at CVCII (UOttawa) and CVCIII (YorkU) for invaluable comments on spoken versions of this paper. Lastly, thanks to my second reader Jack Chambers whose comments in the final leg of this project resulted in a much more readable and convincing final product.

Family: Thanks to my Mom, Dad, and brother Jeff for endless support and encouragement. This paper is dedicated to my Granny, Peggy Rose Denis, whom I lost this winter. I know she’s proud!

Friends: All of my non-linguist friends are very important. Doug, Andy, Davey, Chris, Pat, Weaver, Laura and Ryan let me escape the ivory tower for a while.

Animals: Kismet and Bruce. Meow!

Supervisor: Sali Tagliamonte. I have learned more about being an academic from Sali than anyone ever. To write a paper with Sali is to learn. To have a meeting with Sali is to learn. To have a drink with Sali is to learn. I have worked for and with Sali for five years now and every time we talk I learn more.

Project maker: Becky Roeder. Without Becky agreeing to let me look at the Thunder Bay corpus, this paper would end with the acknowledgments.

Alex: From syntax to soccer; from Gödel to ghosts; from Wittgenstein and Popper to Mulder and Scully - words can’t describe your impact on my work and more so on my life. Thank you for being you and being with me! Without your love and your brilliance surrounding me all year, I’d be lost!
For Peggy Rose Denis (1928-2009)
Table of Contents

Section 1. Introduction ........................................................................................................ 1
Section 2. General methodology: The variationist framework ........................................ 3
  The key concept: The linguistic variable ........................................................................ 3
  Additional concepts and tools ......................................................................................... 4
Section 3. Transmission and diffusion ............................................................................. 7
  Transmission and incrementation ................................................................................... 7
  Diffusion of innovation .................................................................................................. 9
  Applying Labov’s distinction to variationist methodology ............................................. 10
Section 4. Thunder Bay: History, settlement, and corpora ............................................. 13
  History of settlement ...................................................................................................... 13
  Industry, transportation and contact with southern Ontario .......................................... 14
  Thunder Bay today ......................................................................................................... 16
  Corpora ......................................................................................................................... 16
  The Thunder Bay Corpus .............................................................................................. 16
  The Toronto English Archive ......................................................................................... 17
  British English Roots Archive ....................................................................................... 17
Section 5. The transmission of rising have to .................................................................. 18
  The variants ................................................................................................................... 18
  Circumscribing the variable context ............................................................................ 20
  Results: Overall distribution ....................................................................................... 21
  Internal factors ............................................................................................................. 25
  Comparative variable rule analysis .............................................................................. 28
  Discussion ..................................................................................................................... 29
Section 6. The diffusion of innovative be like .................................................................. 32
  The variants ................................................................................................................... 32
  Overall distribution ....................................................................................................... 33
  The source of diffusion ................................................................................................. 36
  Circumscribing the variable context ............................................................................ 40
  Internal factors ............................................................................................................. 40
  Distributional results ..................................................................................................... 42
  Comparative variable rule analysis .............................................................................. 43
Section 7. Conclusions ...................................................................................................... 52
References ........................................................................................................................ 54
Section 1. Introduction

Two ways that language has been argued to change are through transmission and diffusion. Labov (2007) explores the distinction between these two processes and their divergent consequences by highlighting the spread of the innovative NYC short-a system throughout the east coast and the diffusion of the Northern Cities Shift down the St. Louis corridor. Labov argues that changes due to the transmission of a variable phenomenon are characterized by the maintenance of complex linguistic structures through multiple generations of speakers. On the other hand, when a linguistic innovation is diffused to a speech community, the system that underlies the variable grammar of the adopters will be distinct from that of the agents of change.1 This is because the adopters are by and large adults who lack the language learning ability of children. Therefore, replication of the abstract structure of a variable system is imperfect.

Labov’s paper considered only sound changes. Much in the same spirit as Tagliamonte & D’Arcy’s (2009) exploration of Labov’s (2001) incrementation model, this paper seeks to apply Labov’s model of transmission and diffusion to levels above phonology. Are the patterns associated with transmission and diffusion observed in Labov (2007) the same when considering morphosyntactic variables? And if not, how do they differ? To address this question, I examine two variable morphosyntactic linguistic systems in Canadian English: the deontic modality system and the quotative system. Specifically, I employ the comparative method (Tagliamonte 2002) in examining two Canadian speech communities. The first, the centre of Canadian urban English culture, is Toronto. The Toronto speech community has been extensively examined by Tagliamonte and associates in recent years. The SSHRC funded research project Canadian English entering the 21st century (Tagliamonte 2003-2006) assembled a ‘mega-corpus’ of Toronto English representative of the baseline speech of the largest city in Canada. The second speech community examined here is Thunder Bay Ontario. In the summers of 2006 and 2007, Rebecca Roeder collected a modest sized corpus of sociolinguistic

---

1 This terminology is taken from Rogers (1962), one of the most important works on the diffusion of innovation.
interviews with native Thunder Bay residents (Roeder 2008). These two corpora will be compared and contrasted in light of Labov’s distinction between transmission and diffusion.

This paper is structured as follows. In the next section I discuss the broad methodological approach behind this research - the variationist framework (Poplack & Tagliamonte 2001; Tagliamonte 2006). In §3 I detail the definitions of transmission and diffusion of change as put forth by Labov (2007) in order to apply this distinction within the variationist framework. §4 outlines the settlement history of Thunder Bay and discusses the linguistic implications, especially in relation to the Toronto speech community. Here I also provide an overview of the Thunder Bay Corpus, the Toronto English Archive and other corpora to which I make reference. §5 and §6 present the two case studies. The first discusses the deontic modality system and the second, the quotative system. §7 concludes the paper.

---

2 This corpus was made available to me through the Directions of Change project (Tagliamonte 2007-2010).
Section 2. General methodology: The variationist framework

The key concept: The linguistic variable

The framework of variationist sociolinguistics traces its lineage to the seminal work of Weinreich, Labov, and Herzog (1968) and was further developed by Labov (1963, 1966/1982) and his students since the early sixties. Tagliamonte (2006) provides a comprehensive review of the variationist framework and the developments in the field over the last 40 years. The axiom of this framework is that all languages contain variation, all languages change and all linguistic change requires periods of linguistic variability in which languages exhibit “orderly heterogeneity” (Weinreich et al 1968:100). Concretely, languages have linguistic variables, a concept simply defined as “two or more ways of saying the same thing” (Labov 1972a; 1975:9; Sankoff 1980:55). All languages contain choices of different ways of saying the same thing. Choices between different syntactic constructions, morphological affixes, or phonological forms are normal aspects of language.

Furthermore, variables differ across speech communities and across individuals, roughly analogous to the biological contrast of phylogenesis and ontogenesis. If variable X is used 70 percent of the time in the middle age population of a speech community, this does not mean that 70 percent of the middle aged people use X and 30 percent use Y. Rather, sociolinguistics has observed that (ideally) 100 percent of middle-aged people use variable X 70 percent of the time. In this sense, variation is ontogenetic; it is intraspeaker. On the other hand, each speech community is unique in its set of linguistic variables, variables that can behave in different ways in different communities. This phylogenetic sense of variation is closer to what theoretical linguists have called parametric or cross-linguistic variation – different varieties of language have different systems. Both these types of variation are relevant for this paper, which asks the questions: Ontogenetically, what is variable in speakers’ grammars? Phylogenetically, how is the ontogenetic variation structured in different speech communities?

In this paper, I follow the variationist methodology (e.g. Poplack & Tagliamonte 2001:§5.1; Tagliamonte 2006). “The variationist enterprise is essentially, and foremost,
the study of the interplay between variation, social meaning and the evolution and development of the linguistic system itself” (Tagliamonte 2006:5). I will focus on the analysis of evolution of the variable grammar.

It is assumed that intra-speaker variation is part of a structured and orderly system, that is “orderly heterogeneity” (Weinreich et al. 1968:100). The individuals’ choices are not random but are constrained by “the phonological environment, the syntactic context, the discursive function of the utterance, topic, style, situation, and personal and/or sociodemographic characteristics of the speaker or other participants” (Sankoff 1982:151). These constraints are inherent to the variable grammar. The goal of the variationist framework is to characterize the nature of this system.

Variables can be constrained by a number of linguistic (or internal) and social (or external) factors. Internal constraints are linguistic environments where we notice that certain variant are favoured or disfavoured. Variationists model the effect of constraints through statistical techniques, including most predominantly logistic regression. Most variables have several internal constraints that affect the realization of the variable. Some constraints interact with one another. External constraints involve any social information about a speaker. More than 40 years of sociolinguistics has shown that linguistic variables have observerable social patterns. However, this study concentrates on age (as a diagnostic of change) and speech community.

Additional concepts and tools

Two key concepts of the variationist framework are particularly relevant to the current study. The first is the principle of accountability (Labov 1972a:72): “we will report values for every case where the variable element occurs in the relevant environments as we have defined them.” In essence, to look at only the variants that are innovative or interesting would be a fruitless endeavour because variants are embedded within a broader variable system. Individual variants must be examined not in isolation but with respect to “the largest homogenous class in which all subclasses vary in the same way” (Ibid.). As Tagliamonte (2006:13) observes:

---

3 Note, the technical sense of “accountable” in this paper refers strictly to Labov’s principle of accountability.
An accountable analysis demands of the analyst an exhaustive report for every case in which a variable element occurs out of the total number of environments where the variable element could have occurred, but did not.

The second, and highly related, concept is that of *circumscribing the variable context* (Labov 1969:728). This concept involves systematically defining the “relevant environments” (Labov 1972a:72) through “exploratory maneuvers” (Labov 1969:728):

The study of variation is necessarily quantitative, and quantitative analysis necessarily involves counting. At first glance, counting would seem to be a simple operation, but even the simplest type of counting raises a number of subtle and difficult problems. The final decision as to what to count is actually the solution to the problem in hand; this decision is approached only through a long series of exploratory maneuvers.

These maneuvers are as follows (Tagliamonte 2006:13):

1. Identify the total number of tokens in which a feature varies and exclude tokens where one variant is categorically present or absent.
2. Define the variants that can be reliably identified; set aside those tokens that do not fit these definitions.

The first maneuver can be understood as a step to insure that we are studying variation alone.4 The second maneuver insures that we are not including any token of ambiguous nature. As will be discussed in the two case studies below, these two concepts are particularly important. For example, in the deontic modality system there are a number of contexts in which one variant appears categorically in the speech communities examined and therefore must be excluded.

Labov’s third maneuver is to “identify all the sub-categories which would reasonably be relevant in determining the frequency” of a variant (1969:729). This exploratory maneuver reduces to identifying the linguistic constraints on variation as discussed above. To test the effect of these constraints, variationists turn to a modified logistic regression model called the variable rule analysis (Cedergren & Sankoff 1974). Most recently, this variable rule analysis is implemented through the software package GoldVarb X (Tagliamonte, Sankoff, & Smith 2005). The model is essentially a logistic regression that has been specifically constructed to deal with data that is poorly distributed among cells – the kind of data typical in natural speech (Sankoff 1988).

---

4 Guy (1988) suggests that categorical should be understood as anything above 95 percent or below 5 percent.
next section discusses the particular inferences we can draw from the types of evidence available while using the model.

Another important aspect of the variationist methodology is that because it is accountable to the data, it is replicable. Replicability is essential to comparability. Answers to questions surrounding transmission and diffusion that this paper addresses are trivial without replicability because these questions can only be answered through direct comparison of multiple speech communities. If different methodologies are followed for two different speech communities, we are comparing invidious entities. My research specifically replicates the methodologies of two papers: Tagliamonte & D’Arcy (2007a) on the deontic modality system in the Toronto English Archive; and, Tagliamonte & D’Arcy (2007b) on the quotative system in the Toronto English Archive. Each of these will be discussed in the respective sections on deontic modality and quotatives in Thunder Bay.

This section has outlined the general methodological approach of the case studies. The next section will examine the definitions of transmission and diffusion of change more directly and discuss how to apply Labov’s (2007) distinction between the two types of change within a variationist sociolinguistic methodology that appeals to a variable rule analysis.

---

5 That is to say, it is only replicable if all the methodology is described and followed.
Section 3. Transmission and diffusion

Labov (2007) explores the distinction between the transmission and diffusion of linguistic change in great depth. The first is a process that happens during acquisition. In other words, transmission is a parent-to-child process. Diffusion, on the other hand, happens through adult-to-adult interaction - through language and dialect contact. Consequently, these two types of change result in differences to the structure of a variable through time. The general idea discussed by Labov was earlier observed by Kerswill (1996:199-200):

Change induced by dialect and language contact will normally involve the simplification of irregularities, because they are difficult to acquire. […] On the other hand, more gradual (or less traumatic) change within a stable community may involve the preservation or even elaboration of complex features, because children stand a good chance of learning them before the end of their “critical period” of dialect acquisition.
Kerswill (1996:199-200)

Transmission and incrementation

The first process of change this paper is concerned with is the transmission of a change. Transmission is intrinsically linked with another process, incrementation. Labov (2001, esp. Ch. 14) systematically discussed the incrementation process while Tagliamonte & D’Arcy (2009) successfully applied the model to several levels of grammar. Much like children acquires the phonetic inventory, the syntactic parameters and the phonological rules of their caregiver’s language, they also acquires all the features of the parent’s vernacular that are variable (Smith, Durham & Fortune 2009). For example, a feature of my mother’s speech is that she has variable (ing). Sometimes she says, “How’s it going [ng]?” and other times she says, “How’s it goin’ [n]?” In the first case, my mother uses the velar nasal and in the latter the alveolar. Indeed, my speech is the same. Sometimes I use the velar nasal and other times I use the alveolar nasal. With respect to this variable, my mother and I are like every other speaker of English. Of course, this variation in both my vernacular and my mother’s is constrained by a number of linguistic and social factors. The main point is that because this feature is variable for my mom and I learned English from my mother, it is variable for me too.
This is a general principle of language. Stable variation is transmitted in acquisition (Labov 2001:418-421).

What about variants that are not stable, but change over time? Labov (2001:§13.3) argues that these too are transmitted at acquisition. However, in these cases something else happens. Though children first acquire the same level of change as their caregivers, there is a period of what Labov refers to as vernacular reorganization (2001:415).6 Between acquisition and a point of stabilization, observed to be at about the age of 17 (see Labov 2001:ch.14; Tagliamonte & D’Arcy 2009:98), children change the way they use a variable. There may be an increase in frequency of certain variants or changes to the internal organization of variable systems (Labov 2007:346).

For example, my mother might say, “I’m really happy,” or she might say, “I’m very happy.” Rarely would she say, “I’m so happy.” These three intensifying adverbs are used variably in her vernacular. On the other hand, I can say, “I’m so happy” and, “I’m really happy,” but rarely would I say, “I’m very happy.” The argument is that at acquisition I learned my mother’s system, but for more than a decade after, I reorganized my vernacular with respect to the intensifier system, just as everyone in my generation did. Most prominently, we increased our use of really as an intensifying adverb (Tagliamonte 2008).7 Furthermore, how and in what linguistic contexts we used really shifted and spread. My use of this variant increased and the internal organization of my intensifier system changed. Again, I am by no means unique. On the contrary, this is also a generally observed principle of language change. Ongoing change takes place over generations in an incremental pattern (Labov 2001:ch. 14; Tagliamonte & D’Arcy 2009:96). As observed in Labov (2007:346), the internal change to the system follows the incremental pattern whether the change involves “frequency, extent, scope, or specificity of a variable.” Not only is stable variation transmitted through parent-to-child interaction, but so is the variation due to ongoing change. This type of systematic

---

6 This contrasts with proposals in which changes are due to reanalysis of the primary linguistic data at acquisition (Halle 1962; Lightfoot 1989; 1997).
7 Tagliamonte & D’Arcy (2009) find peaks at the age of stabilization across a range of linguistic variables including intensifiers. Uscher (2009) presents a Real-Time study of teenage blog entries exemplifying not only the establishment of dominance of really through adolescence but also the stabilization of the system around 17 years.
incremental generational change is what I will further refer to as change through 

Diffusion of innovation

The second kind of change is the diffusion of innovation. Diffusion occurs through adult-to-adult interaction. Certain variables are straightforwardly diffused across communities. Specifically, lexical items are easily adopted by people well into adulthood (Labov 2001). Again, I will exemplify from personal experience. For no apparent reason in the summer of 2008 I began using the word sick in reference to anything I would have previously referred to as rad, awesome, or cool. After only a couple of weeks, my co-worker picked up the word and still uses it.

If it was so simple as to spend all day with someone for a week to learn their language, the world would be full of polyglots, but that is not the case. Other levels of linguistic structure are not so easily adopted. In fact, Labov (2007:371) observes that the maintenance of “complex patterns of phonetic, grammatical, and lexical specification” requires “an unbroken sequence of parent-to-child transmission.” Furthermore, adults “are less likely to observe and replicate abstract features of language structure” (Labov 2007:369). Tagliamonte & D’Arcy (2007b) do observe communal change such that speakers increased their frequency of the innovative be like through their lifetime (Tagliamonte & D’Arcy 2007b:213). However, these adults “did not exhibit the advanced stages of development with regard to the constraints” (Tagliamonte & D’Arcy 2007b:213), consistent with Labov’s (2007:349) observation that any adult adoption of innovation is characteristically “imperfect.”

What we have are two kinds of changes: change due to transmission, that is, parent-to-child, characterized by the maintenance of internal linguistic constraints on variation; and, change due to diffusion, that is, adult-to-adult, characterized by the loss of internal linguistic constraints on variation. The result of diffusion is that the next generation of speakers who acquire a variable in the adopting speech community will, with respect to the structural conditions, behave differently from the next generation of speakers in the speech community that is the source of the innovation. Hypothetically, if variant X was active in Toronto in Generation 1, the next generation, Generation 2 would preserve (and propagate) the same constraints on realization. At the same time, if Generation 1 from
outside of Toronto was in contact with Generation 1 in Toronto, the adopting speakers might learn certain aspects of the innovative usage, such as the frequency with respect to other variants. However, Generation 1 from outside of Toronto would not acquire the internal structure of variation (i.e. the constraints) as governed in the Toronto speech community. As a result, Generation 2 in the adopting speech community would acquire a different set and/or organization of sociolinguistic constraints from Toronto’s Generation 2. Extrapolating, when linguistic change is diffused from one area to another, the same innovative feature might be observed in both places, but the internal linguistic structure will be different.

Applying Labov’s distinction to variationist methodology

As this paper seeks to apply Labov’s model above the level of phonology, a rigorous methodology is necessary for comparison of speech communities. I turn to the comparative method of variationist sociolinguistics (Tagliamonte 2002). Paramount to this approach are the three lines of evidence for interpreting the results of a variable rule analysis (Poplack & Tagliamonte 2001:92; Tagliamonte 2002:731). These lines of evidence are: 1) statistical significance; 2) hierarchy of constraints; and, 3) relative strength of effect (Poplack & Tagliamonte 2001:Ch.5; Tagliamonte 2006:235).

Statistical significance of a factor group (i.e. Labov’s (1966:729) “sub-categories […] relevant in determining frequencies”) is determined by the statistical model. Does the factor group statistically significantly constrain the realization of a variant? This is determined by a standard .05 p-value.8

The hierarchy of constraints concerns the ordering of factors within a statistically significant factor group. What is the direction of an effect? What contexts favour realization of a variant and what contexts disfavour realization? In variable rule analysis factors with a weight above .5 are said to favour realization while factors below .5 are said to disfavour realization (Tagliamonte 2006:145).

The relative strength of effects concerns the weight of the factor groups. This is determined by the range of factor weights within factor groups. Factor groups exhibiting

8 Though a recent implementation of the variable rule analysis, Rbrul, allows for this threshold to be adjusted. If, for example, five factor groups are being tested, each factor group should be tested with a threshold of 0.01 for an overall error level of 0.05 (Johnson 2009). The Bonferroni Correction is a standard statistical technique.
a large range have more influence on the realization of a variant than factor groups with smaller ranges.

These lines of evidence allow us to assess the hypotheses we are testing. Tagliamonte (2006:235) says:

Similarities and differences in the significance, ordering of constraints and strength of contextual factors provide a microscopic view of the grammar of the data under investigation, from which you may infer the structure of different grammars.

These three lines of evidence have been used to determine the genetic relation between disparate varieties (Poplack & Tagliamonte 1991). If the lines of evidence suggest the same pattern in different varieties, this suggests that these varieties are related. If the varieties are shown to be different with respect to the lines of evidence, “such kinship may be ruled out, at least for the linguistic variable under analysis” (Tagliamonte 2006:241-242).

Applying these methodological notions to Labov’s model, I follow Tagliamonte’s (2006:246) heuristic below:

If the direction of effect of a factor group is shared by varieties, and the effect is not universal, this can be evidence that the varieties have inherited that constraint from a common source. On the other hand, where there are dissimilarities, this can be grounds for concluding that the phenomena in question belong to different linguistic systems.

Two speech communities exhibiting the same variable phenomenon and which have identical variable grammars (as determined by the three lines of evidence) can be interpreted as having a common source. Furthermore, the phenomenon in question has been incrementally transmitted in parallel, generation to generation, in the two speech communities. On the contrary, if two speech communities exhibit the same phenomenon but have different grammars underlying the variation, again determined by the three lines of evidence, this phenomenon can be interpreted as being diffused from one speech community to another speech community, or from somewhere else to both. The different variable grammars indicate that the variable system was not passed on through an unbroken string of adult-to-child acquisition.

Labov’s (2007) model of transmission and diffusion predicts that one variable system known to be historically shared by two speech communities and transmitted in parallel in
both communities will be identical in terms of the three lines of evidence. On the other hand, these three lines of evidence will be different in two communities for a variable system known to be diffused from one community to the other. This paper argues that the ongoing changes to the deontic modality system (particularly the rise of *have to*) in Thunder Bay and Toronto are genetically related phenomena, transmitted in parallel, while changes to the quotative system (particularly the rise of *be like*) were diffused from Toronto to Thunder Bay.
Section 4. Thunder Bay: History, settlement, and corpora

History of settlement

The city of Thunder Bay came into existence with the amalgamation of the former cities of Fort William and Port Arthur, part of Neebing Township and part of Sunian township in January 1970. However, both Fort William and Port Arthur existed as independent urban centres for approximately one hundred years prior (Tronrud & Epp 1995).

The settlement in the late 19th century was not the first time Europeans had lived around the Thunder Bay area. The area was first settled as a post in the fur trade in the 17th century. The French explorer Daniel Greysolon, Sieur du Lhut built Fort Kaministiquia in 1679, and soon after the North West Company established a trading post at the same location. The fur trade did not last and the area was soon abandoned (Stafford 1995:38).

Thunder Bay’s second settlement came as a result of the union of Rupert’s Land9 with the Dominion of Canada in 1870. Like much of settled Northern Ontario, urban development in Thunder Bay quickened with the westward expansion of Canada. One of the terms of western settlement was a transportation system that connected the Pacific to the rest of the country. The railway had a major impact on Thunder Bay. In 1882, it was chosen as the main Canadian Pacific Railway station along Lake Superior. This brought massive amounts of government subsidies to the area, and with the subsidies came migration. Although immigrants from Scandinavia, Ukraine, Italy, Greece and China all arrived at around this time, the Canadian-born, Anglo-Saxon majority “controlled the business and political life” of the city and accounted for well over half of the population (Thunder Bay Multicultural Association 1983:1-2).10 By the turn of the century, Thunder Bay had the second largest population in Northern Ontario after Sault Ste. Marie.

---

9 Rupert’s Land was owned by the Hudson’s Bay Trading Company from 1670 to 1870. The area included all of present day Manitoba, most of Saskatchewan, and parts of Alberta, Nunavut, Ontario, Quebec, Minnesota, Montana and North and South Dakota.
10 Frenette & Jasen (1995:144) report that in the early decades of the city, there was a “clearly defined Anglo-Celtic elite, anxious to confirm its sense of superiority.”
In addition to the railway, Thunder Bay thrived in two other industries: storage of grain in elevators and mining. Grain elevators were designed to store the wheat of the prairies on its way to eastern Canada and Europe. Grain was the major industry in turn-of-the-century Canada. Wheat exports increased nearly 30-fold between 1882 and 1914. Thunder Bay’s strategic location between the wheat of the west and the markets in the east led to significant investment in the city and, ultimately, urban growth. Mining was the third largest industry in the area. Prospecting interest in the area began after the McKellar brothers discovered gold and silver near Thunder Bay in the 1860s. Interest continued throughout the 19th century, enough so that in 1881 mining was the largest single occupation reported in the census (Stafford 1995:42).

With the coming of the railway and the significant economic attraction of two other thriving industries, migration from the south soared. Thunder Bay became closely linked with southern Ontario, particularly Toronto. Frenette & Jasen (1995:146) observe that “Toronto was the centre of civilization to which the local elite looked for inspiration and guidance as they sought to transplant the values and amenities of the British Empire and British Ontario into ‘New Ontario’ [Northwest Ontario] soil.” The elitist attitude led to a partitioning of Canadians of British origin and immigrants in Thunder Bay. As a result, many immigrant communities within Thunder Bay established their own cultural societies and organizations to preserve and promote their ethnic identities.

Besides the railway, transportation to Thunder Bay was limited in its early years. Figure 1 from the Historical Atlas of Canada (Dean et al. 1998) shows the one mode of transportation to Thunder Bay: the railway.

This single transportation link did not last long, though. By 1921, as shown in Figure 2, there were a number of routes in and out of Fort William and Prince Arthur by land and seaway.
Figure 1
Canada’s urban network, 1891

Figure 2
Canada’s urban network 1921
Thunder Bay today

Today, Thunder Bay has a population of approximately 110,000 people and is the most densely populated area in Northern Ontario. Thunder Bay is also home to the only university in Northern Ontario. Lakehead University has over 8,000 undergraduate and post-graduate students.

Corpora

This paper draws from three corpora: The Toronto English Archive (Tagliamonte 2003-2005); the Roots Corpus of British English (Tagliamonte 1996-1998); and the Thunder Bay Corpus (Roeder 2008). The former two are both used as points of comparison with Thunder Bay. All data and observations reported on these two corpora are from previous work by Sali Tagliamonte and her collaborators, unless otherwise stated. The third corpus is The Thunder Bay Corpus (Roeder 2008). This paper is the first to report on variation above the level of phonology in Thunder Bay. All observations of the Thunder Bay deontic modality and quotative systems are new.

The Thunder Bay Corpus

The Thunder Bay Corpus was collected by Rebecca Roeder in the summers of 2006 and 2007. In total, 40 speakers, all born and raised in Thunder Bay were interviewed following standard sociolinguistic interview techniques (Labov 1972b; Tagliamonte 2006).

Table 1 displays the constitution of the Thunder Bay Corpus. The speakers in the corpus are relatively evenly distributed in terms of age and gender.

Table 1

<table>
<thead>
<tr>
<th>Corpus constitution, Thunder Bay Corpus</th>
<th>&lt;29</th>
<th>30-49</th>
<th>&gt;50</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>5</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>TOTAL</td>
<td>17</td>
<td>8</td>
<td>15</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 1 displays the constitution of the Thunder Bay Corpus. The speakers in the corpus are relatively evenly distributed in terms of age and gender.

11 Additionally, as Roeder was interested in the socio-phonetic and phonological aspects of the speech community, a number of word lists were also collected; however, these were not used in the present investigation.
The Toronto English Archive

For all intents and purposes, the Toronto English Archive (TEA) is a “mega-corpus” (Poplack 1989). The TEA was collected between 2003 and 2005 (Tagliamonte 2003-2005; Tagliamonte & D’Arcy 2007b). The 1.5 million-word corpus contains over 200 speakers evenly distributed in terms of age and gender, though a more intensive data collection was used with speakers under 20 due to the focus on incoming change. The speakers were specifically chosen from neighbourhoods in Toronto that represented the highest number of native Torontonians, and, for the most part only those born and raised in Toronto were interviewed (Tagliamonte, Molfenter & King 2004). The goal was a broad picture of “old-line” Toronto English (Tagliamonte 2006b). Speakers range in age from under 10 to over 90. The corpus constitution is shown in Table 2.

Table 2
Corpus constitution Toronto English Archive

<table>
<thead>
<tr>
<th></th>
<th>&lt;29</th>
<th>30-49</th>
<th>&gt;50</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>54</td>
<td>22</td>
<td>26</td>
<td>102</td>
</tr>
<tr>
<td>Female</td>
<td>58</td>
<td>23</td>
<td>35</td>
<td>116</td>
</tr>
<tr>
<td>TOTAL</td>
<td>112</td>
<td>45</td>
<td>61</td>
<td>218</td>
</tr>
</tbody>
</table>

The original token files used in Tagliamonte & D’Arcy (2007a, 2007b) were provided to me by Sali Tagliamonte as a research assistant on the Directions of Change Project and the Toronto English Archive.

British English Roots Archive

The section on deontic modality below will also draw from work by Tagliamonte & Smith (2006). The Roots Archive is composed of interviews with elderly people in eight culturally and geographically isolated communities in England, Scotland, and Northern Ireland. In total, 155 speakers were interviewed between 1997-2000 (Tagliamonte 1996-1998). These speakers represent the oldest living generation of each community.

This Forum paper contrasts and compares the speech of 394 speakers from three different corpora, representing two different major varieties of English.\(^{12}\)

---

\(^{12}\) Only a few of the interviews from the Toronto corpus were conducted by me. I am extremely indebted to the dozens of interviewers, transcribers, project managers, and researchers responsible for the three corpora.
Section 5. The transmission of rising have to

The next two sections discuss two case studies designed to elucidate the linguistic distinctions between change due to transmission and change due to diffusion by considering two changes above the level of phonology. This section examines the deontic modality system in Thunder Bay and compares it to earlier results from Toronto. The deontic modality system has been observed to be in a state of flux since at least early Modern English when the form have to was reanalyzed as a competitor of the modal must (Denison 1993; Krug 2000; Lightfoot 1979). Examples are presented in (1).

(1)  a. What was my favourite job? Oh I have to say the hot-dog business.\textsuperscript{13} (J. Walker/M/59)
    b. I can’t stop until I absolutely must. (P. Johnson/F/40)

This change is further complicated by several additional incoming forms. The variants have got to, got to and need to have entered the system more recently. These are exemplified in (2).

(2)  a. He’s gotta go on the twenty-ninth. He’s got an ear infection. (J. Walker/M/59)
    b. Reef\textsuperscript{14} on the anchor, reef on the fishing rod, reef on the door. You gotta reef it. (T. Wallwork/F/31)
    c. Like, if you want certain jobs, you need to be bilingual. (A. Shaw/M/17)

Tagliamonte & D’Arcy (2007a:82) have observed robust variation in this system in Toronto - particularly the monotonic increase of the variant have to through apparent time. The analysis for Thunder Bay will replicate the analysis carried out in Toronto for comparative purposes. Before addressing the specific methodology, I will discuss the five variants in more detail.

The variants

Brinton (1991:2) suggests that the English modality system in general is undergoing a continuous process of grammaticalization. Today, this development is evident in the

\textsuperscript{13} All examples from Thunder Bay Corpus unless otherwise stated.
\textsuperscript{14} Reef is a lexical item used in Thunder Bay to mean pull.
robust variation between modal and semi-modal forms\(^{15}\) (e.g. *will*/*going to; should/*ought to) (Leech 2003:229). Leech’s (2003) quantitative analysis of the LOB, FLOB, Brown and Frown corpora demonstrates that semi-modals are not drastically displacing full modal forms. In the middle of the last century, semi-modals in British English increased by only 8.5 percent, and in American English by 18.1 percent (Leech 2003:229). However, Leech (2003:230) identifies the exceptional *have to* as the only instance of a semi-modal approaching “the frequency of the corresponding true modal *must*.” Not surprisingly, Tagliamonte & D’Arcy (2007a:71) observe the dramatic decline of *must* in Toronto through apparent time at the expense of rising *have to*, as shown in Table 3.

Table 3
Replication of Tagliamonte & D’Arcy’s (2007:Table 10); Overall distribution of deontic modals in Toronto through apparent time

<table>
<thead>
<tr>
<th></th>
<th>&gt;60</th>
<th>30-59</th>
<th>17-29</th>
<th>9-16</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>have to</em></td>
<td>54%</td>
<td>66%</td>
<td>74%</td>
<td>87%</td>
</tr>
<tr>
<td><em>got to</em></td>
<td>21%</td>
<td>15%</td>
<td>10%</td>
<td>8%</td>
</tr>
<tr>
<td><em>have got to</em></td>
<td>10%</td>
<td>9%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td><em>must</em></td>
<td>7%</td>
<td>3%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td><em>need to</em></td>
<td>8%</td>
<td>13%</td>
<td>10%</td>
<td>3%</td>
</tr>
</tbody>
</table>

This encroachment on the semantic territory occupied by modal *must* has been active since the 15\(^{th}\) century (Tagliamonte & D’Arcy 2007a:50). The diachronic trend of the deontic modality system is summarized in Figure 3 (Tagliamonte & D’Arcy 2007a:50).

Figure 3
Tagliamonte & D’Arcy’s (2007a:Figure 1): Development of deontic modality in English

\(^{15}\) Semi- or quasi-modal verbs are semantically equivalent to modals but differ syntactically “in that they have all the usual properties of other verbs” (Lightfoot 1979:112). One such property is that syntactically full verbs inflect for person and number while modals do not, contrasting *he has to* ... with *he must*... .
There are three other marginal forms in this system: *have got to*, *got to*, and *need to*. Nearly a millennium after *have to* established itself as a variant of deontic modality, *have got to* entered the grammars of English speakers (Krug 2000:53). The earliest examples of *have got to* date to 1860, as in (3). *Got to*, which is suggested to be a reduction of *have got to* (Krug 2000), appears in the mid-twentieth century, as in (4) (Visser 1963-1973:1479).

(3) he always remembers when I’ve *got to* take my doctor’s stuff. (1860, G. Eliot)

(4) ‘What do you tell them that it says?’ -- ‘That they *got to* stop worrying about themselves.’ (1942. J. Cary)

*Need to* has received little attention in the literature despite its rapid recent rise (Leech 2003:229; see also Nokkonen 2006 for a corpora based investigation).  

*Circumscribing the variable context*  
To apply Labov’s model, this paper requires comparability between the new data from Thunder Bay and the Toronto data used by Tagliamonte & D’Arcy’s (2007a:11-12) study of the deontic modality system. The only way to achieve comparability is through exact replication of the circumscription of the variable context. First, all tokens of deontic modality that are potentially variable with the five variants (*have to*, *have got to*, *got to*, *need to*, and *must*) were extracted and coded in GoldVarb X (Tagliamonte, Sankoff, & Smith 2005). Second, I followed the protocols developed in Tagliamonte & D’Arcy (2007a) for exclusion and inclusion of tokens. 

Tagliamonte & D’Arcy (2007a) found a number of contexts to be categorically realized as *have to* in Toronto. Following the principle of accountability (Labov 1972a:72), these contexts were excluded from the analysis. These include tokens of non-present temporal reference and interrogatives. In the Thunder Bay data, non-present temporal reference and interrogatives. In the Thunder Bay data, non-present temporal reference and interrogatives.

---

16 Although Tagliamonte & D’Arcy (2007:51) describe *need to* as likely to be the “most recent layer in the system” (mid 19th century), citing the OED, there is evidence that this construction was productive as early as the 14th century, as in (i) (Visser 1963-1973:1426) and (ii) (OED)

(i) a) *his wisdom nedez every good kyng to haue*  
‘Every good king needs to have this wisdom’ (c1389-1400 Three ME, Sermons (ed. Grisdale) 58, 242)  
b) *we nedte to talk to thim* (1534 St. Th. More)

(ii) A good phisician *nedip to loke wel a-boute and be ful ware.* (c1398 J. Trevisa)
temporal reference was categorically realized as have to. There was only a single example of an interrogative. It was realized as got to. Examples are shown in (5).

(5) a. We had to go across to mother's place to have a bath. (H. Smith/M/88)
   b. I'd say, "How far do I got to walk?" (R. Hall/M/59)

Negative constructions were also excluded because must and the semi-modals do not fall within the same envelope of variation. This is because different variants have different semantic scopes with respect to negation (Brinton 1991:22; Tagliamonte & D’Arcy 2007a:62). With semi-modal forms as in (6), negation negates obligation, leaving open the possibility that the speaker may go to the family Christmas party but is not obligated to do so. However, the full modal form must, in negative contexts negates the verb. In example (7) the speaker is obligated to not attend the Christmas party. These subtle differences entail that negative constructions are not to be included within the variable context.

(6) I don't have to go to the family Christmas party. (D. Shooter/M/14)

(7) I mustn’t go to the family Christmas party.

As with the Toronto data, negatives categorically appear with do-support in the Thunder Bay corpus and never with forms such as haven’t to or mustn’t.

Results: Overall distribution

Having replicated the extraction protocol of Tagliamonte & D’Arcy (2007a) for the Thunder Bay data, I now turn to the results.

Figure 4 above shows the overall distribution of the five variants of deontic modality, must, have to, have got to, got to and need to in three varieties of English: Toronto (Tagliamonte & D’Arcy 2007a), Thunder Bay and rural British English (Tagliamonte & Smith 2006).

Immediately apparent is the parallel distribution of forms in Toronto and Thunder Bay, which contrasts with the British distribution. The parallel distribution is fundamental to the assumption above that Thunder Bay and Toronto share the same deontic modality system. In both varieties have to dominates and all other forms are
These marginal forms are identically distributed in the two varieties such that *got to* is second, *need to* is third, *have got to* is fourth and *must* is least frequent.

Table 4
Deontic modality forms through apparent time, Thunder Bay

<table>
<thead>
<tr>
<th></th>
<th>&gt;50</th>
<th>30-50</th>
<th>&lt;30</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>have to</em></td>
<td>72.9</td>
<td>45.1</td>
<td>84.6</td>
</tr>
<tr>
<td><em>got to</em></td>
<td>16.7</td>
<td>35.3</td>
<td>6.4</td>
</tr>
<tr>
<td><em>have got to</em></td>
<td>3.1</td>
<td>9.8</td>
<td>0.0</td>
</tr>
<tr>
<td><em>need to</em></td>
<td>6.2</td>
<td>7.8</td>
<td>9.0</td>
</tr>
<tr>
<td><em>must</em></td>
<td>1.0</td>
<td>2.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The distribution of forms in apparent time in Thunder Bay, shown in Table 4, differs from the distribution in Toronto presented in Table 3 in a couple of respects.\(^{18}\) There is a

---

\(^{17}\) Though *have got to* has been observed to robustly vary with *have to* in British English (Tagliamonte & Smith 2006), Tagliamonte & D’Arcy (2007a:67) find *have got to* to be a marginal form in Toronto English. The diachronic literature offers an explanation for the minor status of *have got to* in Canadian English. It is not until approximately one hundred years after the establishment of Canadian English as a unique variety (Chambers 1991; Dollinger 2008) that Visser (1963/1973:1479) first observes *have got to*. Early Canadian English was characterized by variation between only *have to* and *must* (Dollinger 2008:213) because at the onset of Canadian English, *have got to* was not present in the deontic modality system of English anywhere. After its first attestation, *have got to* was established as a major variant in the British system. However, this change must not have had much effect on the English spoken in Canada.
noticeable dip in frequency for *have to* in the middle age group. As *have to* declines in this age group, the form *got to* increases. Though this initially looks like the classic U-shape of age grading, comparable to Figure 5 from Chambers & Trudgill (1998:78), there is more to the story.

**Figure 5**  
Classic age grading, Chambers & Trudgill 1998:78

---

Figure 5 demonstrates the phenomenon of age-grading such that the middle age generation has increased usage of the standard form. This particular example is from Trudgill’s (1974) analysis of Norwich English variable (ing) showing the distribution of the variable in four different styles across apparent time. We see a dip in the middle age group in favour of the more socially prestigious velar variant rather than the non-standard alveolar variant.

Despite surface similarity to Figure 5, the data from Thunder Bay in Table 4 is different. Rather than a dip in the middle age group in favour of a socially prestigious form (as in Figure 5), in Thunder Bay the non-standard *got to* is the favoured form. *Got to* has consistently been found to be a non-standard form, from its earliest attestations to

\[18\] When age was run in a logistic regression as the only factor, it was not selected as significant. However, its *p*-value of 0.064 is very close to significance and the effect was in the right direction such that the young group favoured *have to.*
contemporary reports (Tagliamonte & D’Arcy 2007a:79). Indeed, this is not the usual character of age-grading reports because the middle age group favours socially prestigious forms in cases of age-grading. One possibility is that got to has covert prestige in Thunder Bay.

One way to address this issue is to ‘explode’ our data by plotting the individual speakers according to frequency. Do all middle age people use got to more than have to, confirming age grading? Figure 6 displays a scatterplot of individuals in the middle age generation by frequency of got to.

Figure 6
Scatterplot of individuals, got to

This chart suggests that there is more to the story than age grading. Two individuals use got to at an uncharacteristically high rate. These two individuals, one male and one female, are marked with xs in the scatterplot. These speakers were interviewed at the same time, live together, are unemployed, and have suffered individual, work-related accidents. They have no post-secondary education and worked in factory-type jobs before their accidents. It is unsurprising then that got to, found to be a marker of

---

19 The forms have got to and got to have been viewed as “colloquial”, “meaningless” (Visser 1963-1973:1479), and “nonstandard” (Krug 2000:62).
working-class speech in Toronto, dominates their deontic modality system.\textsuperscript{20} At this point, a conclusion of age-grading is highly unlikely. There are not enough speakers in the age groups.

In sum, the overall distribution results suggest that Thunder Bay and Toronto are the same in terms of frequency of the forms used for deontic modality. Both varieties are characterized by a rise of have to while the other forms are marginal. However, it would be premature to conclude that the changes in the deontic modality system in Toronto and Thunder Bay have been transmitted in parallel, rather than diffused, based on frequency alone. As Labov (2001:447) observes, the speech of adults is not stable through the lifetime. Changes in frequency have been reported. For example, Tagliamonte and D’Arcy (2007b:213) find that the frequency of the quotative be like has increased through real time in Canadian English. In 1995, speakers in their late teens used be like 13 percent of the time. Approximately a decade later, speakers in the same generation (now in their late twenties and early thirties) were observed to use be like between 31 and 58 percent indicating that speakers changed their rates of usage well beyond the point of stabilization. It is possible that speakers in Thunder Bay have adjusted their frequencies of have to to match the Toronto community. Thus, the distributional findings reported above are not sufficient evidence for transmission.

Fortunately, Tagliamonte & D’Arcy (2007b:213) also observe that the underlying constraints on be like did not change for older speakers in Toronto. Rather, the system remained incipient. Tagliamonte & D’Arcy (2007b:213) conclude that “frequencies of linguistic forms are labile […] but the grammar underlying them is not.” Therefore, to address the question of transmission within the Thunder Bay deontic modality system, the present analysis must go beyond frequency distributions and examine the underlying system. If Toronto and Thunder Bay’s variable grammars are identical this suggests that the systems are one and the same, transmitted in parallel from generation to generation.

\textit{Internal factors}

In order to examine the internal grammars underlying the deontic modality system in Thunder Bay, we must examine the internal constraints active in the system. Following

\textsuperscript{20}The third speaker, age 59, whose rate of got to is comparable to the two individuals discussed here is also blue collar. He is self-proclaimed to have “not much” education and he works as a mechanic.
Tagliamonte & D’Arcy (2007a:64-65), three factors were coded: Grammatical Person of the Subject, Definiteness of the Subject, and Source of Obligation. The first two of these factor groups were the result of operationalizing the claim that different variants exhibit different levels of obligation. Coates (1983:32) has claimed that Grammatical Person of the Subject influences the choice of forms. 2nd person subject, as in (8), exhibit the highest degree of obligation; tokens with 3rd person subjects, as in (9), exhibit the lowest degree of obligation; and, tokens with 1st person subjects as in (10) are somewhere in the middle.

(8) I can't skip practice right. If one day, "Oh yeah you have to come to practice." (A. Stone/M/14)
(9) he's gotta go on the twenty-ninth he's got ear infection [sic] like the bone here is all infected. (J. Walker/M/59)
(10) What was my favourite job? I must say, hot dog vendor. (J. Walker/M/59)

However, Tagliamonte & Smith (2006:359) and Tagliamonte & D’Arcy (2007a:65) have observed that the Grammatical Person of the Subject factor group highly interacts with the Definiteness of the Subject. Both 2nd person and 3rd person subjects can be either definite as in (11) or generic as in (12). Definiteness was also coded.

(11) Oh yeah, yeah. And there's- I see people who are burnt-out who just- they need to retire. (S. Black/F/53)
(12) I-mean obviously they have to have some of their hard industry there but- you ever try and find a hotel room in Duluth, it's pretty tough in the summertime. (J. Schroder/M/63)

The third constraint that has consistently been found to affect realization of deontic modality forms is the Source of Obligation (Tagliamonte & Smith 2006:362; Tagliamonte & D’Arcy 2007a:65). This factor group can be divided into two factors, objective (or external) obligation and subjective (or internal) obligation. These are contrasted in that in tokens of objective obligation, the Source of Obligation is from outside of the speaker, as in (13) where the speaker says that sometimes when he plays golf it starts to rain and since he is already out on the course he has to play in the rain. On the other hand, tokens of subjective obligation have their source from within the speaker, as in (14) where Speaker 2 is saying that he has to stay in shape over the summer in order to be fit for the sports he plays.
27

(13) well you don't start in the rain, but sometimes you have to play in the rain (J. Schroder/M/63)

(14) <Speaker 1> Especially now that since you're playing the sports so much. <Speaker 1> <Speaker 2> Yeah, gotta stay in shape over the summer too. <Speaker 2> (D. Shooter/M/14)

Tagliamonte (2004), Tagliamonte & Smith (2006) and Tagliamonte & D’Arcy (2007a) all report that these three factors interact considerably. Table 5 shows this interaction in the Toronto data. This table shows the total N’s in the cells of cross tabulations of the Source of Obligation by the Grammatical Person in Toronto.

Table 5
Interactions in the system, Toronto

<table>
<thead>
<tr>
<th></th>
<th>1st def.</th>
<th>2nd def.</th>
<th>3rd def.</th>
<th>2nd gen.</th>
<th>3rd gen.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>546</td>
<td>151</td>
</tr>
<tr>
<td>Def. Objective</td>
<td>212</td>
<td>20</td>
<td>111</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Def. Subjective</td>
<td>268</td>
<td>108</td>
<td>19</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Notice that 2nd person definite subjects have a clear preference for subjective readings, 3rd person definite subjects have a clear preference for objective readings, and 1st persons show a split. Note also that generics, which are inherently objective, must be separated out. Leaving them in would skew the distribution of forms for the objective tokens in whatever direction the generics favoured.

Table 6 below shows that the same distribution obtains in Thunder Bay.

Table 6
Interactions in the system, Thunder Bay

<table>
<thead>
<tr>
<th></th>
<th>1st def.</th>
<th>2nd def.</th>
<th>3rd def.</th>
<th>2nd gen.</th>
<th>3rd gen.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>4</td>
</tr>
<tr>
<td>Def. Objective</td>
<td>60</td>
<td>5</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Def. Subjective</td>
<td>35</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Tagliamonte & D’Arcy (2007a:66) resolved the interaction by collapsing the factor groups of Grammatical Person, Definiteness, and Source of Obligation into one interaction group entitled type of subject. This paper follows their methodology. The only linguistic constraint tested for was the type of subject with the factors, generic, definite/objective, and definite/subjective (Tagliamonte & Smith 2006:27; Tagliamonte & D’Arcy 2007a:78).

Tagliamonte and Smith (2006) and Tagliamonte & D’Arcy (2007a) observe a different distribution of variants by these three factors in British English and Toronto.
English. In British English, *have to* is favoured by definite objective subjects, *must* is favoured by definite subjective subjects, and *(have) got to* is favoured by generic subjects. On the other hand, in Toronto English *have to* is favoured by both definite objective subjects and generic subjects while *got to* is favoured by definite subjective subjects. These differences are summarized in Table 7.

Table 7  
Deontic forms by reference and pragmatics, British & Toronto English

<table>
<thead>
<tr>
<th>Type of subject</th>
<th>BrE</th>
<th>TorE</th>
<th>BrE</th>
<th>TorE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Def. Subjective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Def. Objective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Comparative variable rule analysis*

As mentioned earlier, this same partitioning of form by function can be achieved in Thunder Bay by subjecting the data to an identical variable rule analysis. This allows for accountable comparison of the speech communities, which is necessary for addressing the question of transmission or diffusion.

The variable rule analysis of *have to*, testing the effect of type of subject is presented in Table 8. The data is separated into the Thunder Bay data and the Toronto data (Tagliamonte & D’Arcy 2007a). Recall that the three lines of sociolinguistic evidence and the comparative method can be used to see how linguistic changes are spreading.

The variable rule analysis of the Thunder Bay data replicated those of Tagliamonte & D’Arcy (2007a) in Toronto.

Table 8  
Variable rule analysis results, *have to*

| Type of subject | Thunder Bay |  |  | Toronto |  |  |
|-----------------|-------------|-------------|-------------|-------------|-------------|
| Input           | .751        | .699        |  |  |  |  |
| Total N         | 225         | 1627        |  |  |  |  |
| Type of subject | FW % N      | FW % N      |  |  |  |  |
| Generic         | .58 75.9 79 | .56 78.3 727|  |  |  |  |
| Def. Obj.       | .60 78.4 88 | .61 83.9 373|  |  |  |  |
| Def. Sub.       | .23 53.1 49 | .30 55.6 401|  |  |  |  |
| Range           | 37          | 30          |  |  |  |  |

The frequency distributions alone are quite telling. Thunder Bay and Toronto appear to share the same hierarchical effect such that *have to* is most common in definite
objective subject types followed closely by generic subjects. This evidence suggests that Labov’s characterization of transmission, which he based on phonological studies, will also be applicable to morphosyntax. The distribution of have to by type of subject is the same in the two speech communities.

This result is further strengthened by the variable rule analysis. The factor group type of subject was found to be statistically significant in both Toronto and Thunder Bay for determining the realization of have to. The relative strength of the factor groups is not applicable to this analysis since only a single interaction group was tested for. However, the constraint hierarchy provides important evidence. It is identical in Thunder Bay and Toronto. Have to is most favoured in tokens of definite objective context and is favoured also in tokens of generic context; this contrasts with what Tagliamonte & Smith (2006) report for the British data. Lastly, there is a strong disfavouring effect of have to in definite subjective contexts.

Discussion
The comparative method supports an interpretation of parallel transmission of the rise of have to in Toronto and Thunder Bay consistent with Labov’s (2007) model of transmission and diffusion. Labov (2007) observes that in cases of transmission of phonological features, constraints are propagated. In comparative analysis of Thunder Bay and Toronto, the constraints on the deontic modality system are propagated. I observe that the deontic modality systems in Thunder Bay and Toronto are the same systems and, extrapolating from Labov (2007), I propose that they are the same by virtue of having been transmitted in parallel since the settlement of Thunder Bay. The results discussed in this section provide an interesting extension to the findings of Tagliamonte & D’Arcy (2007a). While Tagliamonte & D’Arcy (2007a:82) note that Toronto English “distinguishes itself by a unique profile in the grammatical constraints operating on this system,” this paper suggests that the uniqueness of Toronto English can be extended to Ontario more generally. This is consistent with the observation that urban, middle-class Canadian English is one monolithic homogenous speech community from the Ontario-Quebec border to the Pacific Ocean (Priestley 1951; Chambers & Hardwick 1986; Chambers 1991).
The historical record provides an explanation. As §4 explained, Thunder Bay was first permanently settled by Europeans around 1885. More than 50 percent of the residents of the new city were Canadian born, Anglo-Saxon migrants from southern Ontario (Frenette & Jasen 1995). Southern Ontario English, as spoken by these 19th century settlers, was the foundation of Thunder Bay English.

Based on his Corpus of Early Ontario Correspondences\(^2\), Dollinger (2008:213) reports that only \textit{have to} and \textit{must} were present as markers of deontic modality in early Canadian English. Consistent with the trajectory of change reported in Tagliamonte & D’Arcy (2007a), Dollinger reports a rise of \textit{have to} by 22 percent between 1776-1799 and 1800-1849.

Figure 7
An amalgamation of Tagliamonte & D’Arcy (2007a) [left] and Dollinger (2008) [right]

The historical and contemporary findings can be merged to reveal the full historical trajectory of the rise of \textit{have to} in Canadian English. This amalgamation of results is presented in Figure 7.

Unlike the results reported for British English in Tagliamonte & Smith (2006:354), in which \textit{have got to} robustly varies with \textit{have to} and \textit{must}, Toronto English is dominated by \textit{have to}. The historical literature on Thunder Bay’s settlement history suggests a heavy migration from southern Ontario and with it, the strong influence of Toronto. Because of

\(^2\) The Corpus of Early Ontario Correspondences is a compilation of pre-confederate diary entries, letters and newspapers written by early Ontarians. The corpus is divided into two time periods, 1776-1799 and 1800-1849.
Thunder Bay’s settlement from southern Ontario, the trajectory of change underway there in the early 1800s, shown on the right side of Figure 7, was transplanted to Thunder Bay at the time of settlement. Following Zelinsky’s (1992) Doctrine of First Effective Settlement, which has been applied to linguistics (Labov 2007 inter alia), the grammar of the first effective settlement will continue unaffected by future immigration.22

Whenever an empty territory undergoes settlement, or an earlier population is dislodged by invaders, the specific characteristics of the first group able to effect a viable, self-perpetuating society are of crucial significance to the later social and cultural geography of the area, no matter how tiny the initial band of settlers may have been23 … in terms of lasting impact, the activities of a few hundred, or even a few score, initial colonizers can mean much more for the cultural geography of a place than the contributions of tens of thousands of new immigrants generations later. (Zelinsky 1992:13-14)

When a community is settled, the culture, including the language, of the first settlers will continue on for generations to come. The specific assumption then is that today’s Thunder Bay speech community is a direct descendent of the dialect spoken by the first settlers. This assumption is borne out in the linguistic data presented above. For the deontic modality system, variable rule analyses and Labov’s model of transmission reveal a diachronically shared system by the Thunder Bay and Toronto speech communities.

The rise of have to in the deontic modality system of this larger speech community has been incremented as “a faithfully reproduced pattern” generation by generation in both communities. The rise of have to, a change that was active in southern Ontario at the onset of Canadian English in the 18th century (Dollinger 2008), was transplanted to Thunder Bay at the time of the city’s settlement in the late 19th century, and the same change in progress has been transmitted in parallel in both communities.

---

22 Mufwene (2001:29) similarly discusses the Founder’s Principle, a concept from evolutionary biology, “to explain how structural features of creoles have been predetermined to a large extent (though not exclusively) by characteristics of the vernaculars spoken by the populations that founded the colonies in which they developed.”

23 Naomi Nagy and Alexandra Motut (pc.) point out the apparent tautological nature of this definition. Essentially, Zelinsky is saying the first effective settlement will be the first effective settlement. How does one define a “viable, self-perpetuating society” without reference to the “later social and cultural geography of the area”? Zelinsky does not address the distinction explicit. An exhaustive search of the cultural geography literature has come up with no mention of this cyclic argument. For now I will assume that Zelinsky must define a settlement in a physical sense, not a cultural sense. This divorces the two propositions: effective physical settlement leads to effective cultural settlement.
Section 6. The diffusion of innovative *be like*

This section explores diffusion, the second aspect of Labov’s (2007) model, by examining the innovative quotative *be like* in Thunder Bay. *Be like* is currently encroaching on the quotative system of contemporary Englishes around the world (Blyth et al. 1990; Ferrara & Bell 1995 (American English); Tagliamonte & Hudson 1999 (Canadian and British English; D’Arcy 2004 (Newfoundland English); Tagliamonte & D’Arcy 2007b (Toronto English); Buchstaller & D’Arcy fc (American, British, New Zealand English)). Tagliamonte & D’Arcy (2007b:199) consider the rise of *be like* to be a “vigorous change” above the level of phonology. Between 1995 (Tagliamonte & Hudson 1999) and 2002 (Tagliamonte & D’Arcy 2004:501), *be like* increased in frequency from 13 percent of all quotatives to a striking 58 percent in Toronto.

As in the previous section, Tagliamonte & D’Arcy’s (2007b) large scale study of the Toronto English Archive is the point of comparison for the Thunder Bay data. Following the methodology of §3, we can compare and contrast the rise of *be like* in the quotative systems in Toronto and Thunder Bay: are the patterns of usage observed in Toronto the same in the Thunder Bay data set? Are the constraints active in Toronto’s system active in Thunder Bay’s? Whatever the answer to these questions, can we explain the results within Labov’s model? Since *be like* could only have been diffused to Thunder Bay and assuming that Labov’s model is accurate then this paper anticipates that its well known constraints may not be present in Thunder Bay.

*The variants*

The quotative system in Canadian English is characterized by a larger number of variants than the deontic modality system. Most common are the variants *say, think, go, be like* and the *null* variant, but there are also a number of marginal variants that complete the system, including *tell, scream, write, figure, ask, mention, decide, yell, wonder, call* and *explain*, among others, making up 4.7 percent of the Toronto system (Tagliamonte & D’Arcy 2007b:217). The focus of this section will be on the nature of *be like* with respect to all the variants in the quotative system.
Be like was first reported in print in Butters’ (1982:149) editor’s note in American Speech. The first systematic study of the variant was Blyth, Rechtenwald, & Wang (1990). Attitudinal results reported by Blyth et al. (1990:224) found that be like was highly associated with the stereotypical California Valley Girl – the embodiment of 1980s affluent, self-absorbed, upper-middle class girls living in the San Fernando Valley. Despite this stereotype, males in the study were found to be more likely to use be like than females. The gender effect has been reported inconsistently since. Romaine and Lange (1991:228) report a female preference, and Ferrara and Bell (1995:273) similarly report that females are leading; however, as the change progressed, the effect was neutralized. Tagliamonte and Hudson (1999:160) report non-significance in Canadian English from the mid-nineties though Tagliamonte & D’Arcy (2007b:208) show that the female preference is active in Toronto, but only among the youngest speakers. Whether or not be like can be affirmed as a feature of female speech is inconclusive; however, that be like originated in California in the 1980s seems to be accurate. We know that before the 1980s, be like was never used in English but has since spread to varieties in Canada, the UK, and Oceania.24

Overall distribution
Table 9 shows the overall distribution of the top 5 quotatives (be like, say, think, go, null) and all other variants in the speech of the youngest generation in Thunder Bay and Toronto (as reported in Tagliamonte & D’Arcy 2007b).

Table 9
Overall distribution of quotative forms, Thunder Bay & Toronto

<table>
<thead>
<tr>
<th></th>
<th>be like</th>
<th>say</th>
<th>go</th>
<th>think</th>
<th>null</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Thunder Bay</td>
<td>81.1</td>
<td>159</td>
<td>9.2</td>
<td>18</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>Toronto</td>
<td>63.7</td>
<td>2093</td>
<td>13.3</td>
<td>436</td>
<td>3.4</td>
<td>112</td>
</tr>
</tbody>
</table>

The data from Thunder Bay has nearly 20 percent more be like tokens than Toronto. This preference for be like in Thunder Bay is at the expense of the diversity of forms used. Say, go, think, the null variant and all other forms are used less in Thunder Bay than in Toronto. However, although Toronto and Thunder Bay appear to differ in terms

---

24 Parallels in German and Swedish have also been reported (Golato 2000).
of the overall distribution of forms Figure 8 suggests that they are more similar to one another than to other varieties that have been reported in the literature.

Figure 8
Distribution of quotative forms, Thunder Bay, Toronto, Britain, Ottawa

The distribution pattern of the top five variants is nearly identical in Toronto and Thunder Bay; much more so than the British data (York, Tagliamonte & Hudson 1999) or Ottawa (Tagliamonte & Hudson 1999). Figure 8 also reveals the general cline of development of be like through time. Beginning with the earliest data from Ottawa and ending with the latest data from Thunder Bay, there is a monotonic increase of be like. Ottawa has less than 15 percent be like and the British data has just a little bit more. Fast forward nearly ten years and Toronto has a rate of over 60 percent. By 2007, Thunder Bay’s rate is over 80 percent. The 20 percent difference between Toronto and Thunder Bay for be like observed in Table 9 may simply represent the next stage of development of be like in English. Indeed, the Thunder Bay speech community fits with the general spread of be like since the 1980s. Figure 9 below displays a view of be like through apparent time in Thunder Bay, ‘exploded’ by individual. The scatterplot’s X-axis displays age and the Y-axis shows frequency of be like. Each marker represents one speaker.
Be like is completely absent in the oldest speakers in Thunder Bay, consistent with be like’s origin in the 1980s. No one over the age of 40 (over the age of about 20 in the 80s) would have acquired the feature. The feature could not have been transmitted from the older generation to the younger generation. It must have arrived in Thunder Bay via diffusion.

Figure 9
The trajectory of be like through apparent time in Thunder Bay (Scatterplot)

The oldest user of be like is 40 years old - a teenager in the 80s. The rate of be like for this speaker is also the lowest of those who have the variant at only 4 percent of 55 tokens. This perspective of the Thunder Bay data is particularly useful for deciding what the appropriate way is to group speakers in terms of age.25 Figure 9 clearly demonstrates three groups in the data. There are those over 40 who do not have be like, those speakers who are between 30 and 40 who, with the exception of one speaker, have marginal rates of be like, and those in their teens, for whom be like is robust. I can now make a reasonable comparison with the Toronto data by comparing Thunder Bay and Toronto in

---

25 To add a methodological point, viewing individual speakers like this is much more informative than the to ‘bin’ speakers. This is especially true for small data sets (Tagliamonte pc.).
The oldest speakers in both communities do not have *be like*. The middle age speakers have a rate of *be like* around 15 percent. Note that this is the rate that Tagliamonte & Hudson (1999) report for teens in another Canadian speech community (Ottawa) in 1995. These speakers would now be in this 30-40 age range. In both communities, the quotative system in the youngest generation is dominated by *be like*.

Figures 9 and 10 offer an important point regarding Labov’s transmission and diffusion. *Be like* is absent in the oldest generation, jibing with the fact that the innovative form was diffused to the two communities. Recall that the empirical evidence for diffusion comes from the three lines of sociolinguistic evidence. A variable grammar underlying a system known to be diffused from one community to the other will be observably different in the two communities. Since this paper adopts Labov’s distinction between transmission and diffusion, it is necessary to establish Toronto as Thunder Bay’s source of *be like*. Otherwise the comparison is moot.

*The source of diffusion*

Two pieces of demographic evidence support the argument that Toronto is Thunder Bay’s source of *be like*. The first comes from the gravity model of diffusion (Trudgill
This model predicts that the influence of one speech community on another is a function of the distance between and population of the speech communities. The gravity model predicts a balance between distance and population. Additionally, linguistic factors such as prior existing linguistic similarity and the type of linguistic change, and geographic factors such as physical barriers (lakes, mountains), political frontiers, and infrastructure (roads, communication etc.) all have an effect on the strength of influence of one speech community on another.

Thunder Bay is the only metropolis in Northern Ontario. The closest large city is Winnipeg in Manitoba to the west. Duluth, Minnesota is the second closest, though the two communities are separated by the American-Canadian border. Without a doubt, the speech community that should have the largest influence on the speech of Thunder Bay is Toronto. As we have seen, The historical literature indicates that Toronto was the cultural and social model for the developing town at the turn of the century, and certainly Toronto continues to be the largest urban centre in Canada.

We can approximate Toronto’s influence on Thunder Bay through time by considering the gravity model equation of influence (Trudgill 1974b:235). The equation is presented in (15).

\[
I_{ij} = s \cdot \frac{P_i P_j}{(d_{ij})^2} \cdot \frac{P_i}{P_i + P_j}
\]

\(I_{ij}\) is the influence of community \(i\) on community \(j\), \(P\) is the population of the community and \(d\) is the distance between these communities. The variable \(s\) expresses linguistic similarity between the communities. As a preliminary exercise, this formula was applied to six different communities surrounding Thunder Bay at two periods in

---

26 A second model of diffusion is the cascade model (see Labov 2003). The crucial difference between these models is the lack of importance of physical distance between two speech communities in the cascade model. Both models have been found to make solid predictions: Trudgill’s (1974b) foundational study of the Brunlanes region of Norway and the East Anglia region of England found promising results for the gravity model; Labov (2003) reports on studies in which the cascade model was more effective.

27 Chambers (2009) argues that political border towns tend to be bastions for certain features associated with the territorial variety.

28 \(s\) can be calculated based on a number of parameters and levels (see Trudgill 1974b). For the present case, \(s\) was ignored because all of the dialects belong to a North American variety and are similar at some level.
time, returning the influence of each town through the city’s history. The results are presented in Figure 11.

Figure 11
Influences on Thunder Bay

![Influence bar chart]

Distance was determined by shortest road distance between Thunder Bay and the six communities as determined by Google Maps. Population statistics at each period were taken from various Statistics Canada reports and from an equivalent American source for Duluth. The 1980s population is from the 1986 census. Population figures for Sault Ste. Marie were unavailable for 1980s.

Figure 11 shows that Toronto’s influence on Thunder Bay is greater than any of the other urban centres, suggesting a likely source of linguistic innovation. In the 1980s, when *be like* began to spread, Toronto’s influence as determined by the gravity model was stronger than any other community. This influence continues today, supporting the argument that Toronto is the likely source of *be like* for Thunder Bay.

---

The demographics of the student population of Lakehead University add further evidence that Toronto, or at least the southern Ontario region, is a likely candidate for the source of be like to Thunder Bay. Figure 12 shows the location of all universities in Ontario. The sole marker on this map in northern Ontario is Lakehead University. Lakehead’s population is primarily composed of students from Northwestern Ontario. They make up the majority 50.8 percent. Out-of-province students comprise 5.7 percent and international students make up 2.3 percent. The remaining 41.2 percent of the student body is from southern Ontario. Lakehead’s current population of approximately 10 000 university aged people and Lakehead’s 7300 students entail that approximately 3000 young people living in Thunder Bay (30 percent) are from southern Ontario. These speakers are likely agents of change – the diffusers of be like.

Having established that Toronto is the most likely source of the innovative be like quotative to Thunder Bay, Labov’s model of diffusion and the predictions regarding the three lines of evidence discussed in §3 can be applied to our comparison of Toronto and Thunder Bay.
Circumscribing the variable context

As in the previous section, the first step in comparative sociolinguistics is to accurately circumscribe the variable context in the two communities. Because this paper compares Thunder Bay to Toronto, I replicated the variable context as detailed in Tagliamonte & D’Arcy (2007b). To begin with, every instance of “a verb introducing constructed dialogue” was extracted (Tagliamonte & Hudson 1999:154). Tokens of *think* and *say* that introduce indirect speech, as in (16), were categorically excluded.

(16) He said he was tired.

These tokens were not extracted because quotative forms, such as *be like*, are not possible in these contexts. Furthermore, two environments were excluded due to different kinds of ambiguity. Incomplete and incomprehensible tokens were excluded. These could not be reliably coded for a number of factor groups.

Internal factors

Three linguistic factors have been found to constrain the realization of *be like* in Toronto: Grammatical Person of the Subject; Content of the Quote; and, Tense/Morphology.

Grammatical Person of the Subject is a straightforward factor group: all tokens were coded for grammatical person of the subject. Tagliamonte & D’Arcy (2007b: 209) report a favouring effect for 1st person subjects in all three age groups that use *be like*. However, Tagliamonte & D’Arcy (2007b:208) suggest that this effect may be residual of the pragmatic effect of Content of the Quote.

Quotative verbs can introduce a number of different types of content. Most common is Direct Speech, as in (17), where a speaker is directly quoting or paraphrasing a speech event.

(17) And then I see the grass move, and I'm like, "Princess, get over here!" So she turns around and she comes back, like normally she doesn't. And the skunk started following her! (M. McGuinness/M/40)

A second type of Content of the Quote is Inner Dialogue, as in (18). In these tokens, speakers are reporting on thoughts.

(18) It's very confusing and I'm like, "Why am I learning this?" (K. Howard/F/19)
Early on, *be like* was reported to have originated as a marker of internal speech (Butters 1982; Tannen 1986). Tagliamonte & D’Arcy (2007b:208) observe that this constraint has considerably weakened in their apparent time data. In the middle aged generation, this factor highly constrains the application of *be like*, but in the youngest generation the factor group is the lowest. This weakening has been argued (D’Arcy 2004:335) to be an indicator of grammaticalization, so that *be like* came in as a marker of internal dialogue and has since spread to introduce all kinds of quotable content, including Direct Speech, but also instances of non-lexical content, hypotheticals and written language.

The third factor group considered by Tagliamonte & D’Arcy (2007b) is the temporal reference and surface morphology of the quotative verb. *Be like* has been observed to appear in historical present (HP) contexts (Singler 2001). The historical present is a “specific and highly circumscribed construction in which surface morphology is present tense but the context of temporal reference is past tense,” as in (19) (Tagliamonte & D’Arcy 2007b:204).

(19) And she was[^PAST] like really having a bad day, she looked[^PAST] at me and she's[^PAST] like, "No!" (A. O’Conner/F/14)

This context must be contrasted with other tense types. Following Tagliamonte & D’Arcy (2007b:204), my analysis also makes a three-way distinction between HP, unambiguous present tense/morphology (20) and past tense (21).

(20) I want to talk on M-S-N and then Kobe's like, "Come on, I gotta go play my game, blah blah blah." (A. O’Conner/F/14)
(21) I remember when the writers strike started just earlier this week, all my friends were like, "What's happening with the writers strike?" (I. Calder/M/20)

Furthermore, a number of other tense/morphology contexts were coded for including habitual *would* (22), infinitives (23), progressives (24), and modals (25), among others.

(22) My parents would always- about every three or four years we-they'd say, "Let's go," and so off we'd go. (S. Black/F/53)
(23) It's not even rules it's just they're not used to living with other people. Which makes it harder for them to *be like*, "okay." (K. Howard/F/19)
(24) I heard something upstairs so I went up to see and then I saw my mom and dad putting down presents saying, "From Santa." (D. O’Conner/M/12)

(25) You're not just listening to whatever they're playing. You can decide, "Oh I want country," or, "No, I don't want country, I want this kind." (P. Johnson/F/40)

**Distributional results**

For an overall comparison of the two communities, consider some distributional data, first examining the distribution of *be like* by collocating tense/morphology. There is roughly the same breadth of distribution of *be like* tokens by tense in Thunder Bay and Toronto as Figure 13 shows.

Figure 13
Distribution of *be like* tokens by tense/morphology

This suggests two similar systems with respect to the Tense/Morphology factor group though there is one notable difference. *Be like* is used approximately 15 percent less in the simple past tense in Thunder Bay than Toronto while the form is used 10 percent more in the simple present.
There is a second nuance in the Thunder Bay system, this time with respect to the Content of the Quote constraint as shown in Table 10.

In Toronto, 73 percent of the quotative tokens are tokens of Direct Speech, 18 percent are Inner Dialogue, and the remaining tokens represent other types of content such as written material and non-lexicalized sounds. Thunder Bay exhibits a slightly different profile. The vast majority of tokens (80 percent) are Direct Speech while only 9 percent are Inner Dialogue. This is half the number of tokens of Inner Dialogue than in Toronto. The remaining 11 percent of tokens in Thunder Bay are represented by other types of content. I now turn to a comparative variable rule analysis and application of Labov’s model.

**Comparative variable rule analysis**

To address the question of diffusion, two identical variable rule analyses were run on the Toronto data and Thunder Bay data. These analyses were composed of independent runs on the middle age generations (30 to 40) and the young generations (10 to 20). The factor groups tested for are those discussed above: Tense/Morphology, Content of the Quote, Grammatical Person and Sex.

Table 11 presents the result of the variable rule analysis for Toronto. The results are from the identical source data as Tagliamonte & D’Arcy (2007b) but new GoldVarb runs were performed for direct comparability with Thunder Bay.

The strongest factor constraining realization of be like is the Tense/Morphology group. Within this group, HP highly favours be like and past tense disfavours be like. Through apparent time, present temporal reference with present tense morphology has become less favoured. Tagliamonte & D’Arcy (2007b:209) argue that be like began with an association to the present tense generally, but through apparent time, HP is singularly favoured. Content of the Quote has a favouring effect for Inner Dialogue in both generations, though this effect is weakened in apparent time from having a range of 27 in the 30-40 age group to a range of 12 in the 10-20 age group.
Tagliamonte & D’Arcy (2007b:212) argue that though *be like* arrived in Toronto and filled the grammatical niche of “speaker thought process,” it began to specialize in a different way, to that of a narrative marker “by its initial association with monologue.” The HP factor became much more important and the Content of the Quote factor weakened. Lastly, the person constraint is weak in both generations, as reported by Tagliamonte & D’Arcy (2007b). The Sex effect is as reported by Tagliamonte & D’Arcy. *Be like* arrives in the community without social status but in the next generation an association with female speakers develops.

The real question concerns *be like* in Thunder Bay. Are the same effects active in Thunder Bay as in Toronto, or are they different, as the evidence thus far presented and applied to Labov’s model would predict?
Table 12 presents the results of an identical analysis of Thunder Bay as shown for Toronto in Table 11.

There are a number of important points in Table 12 to contrast with the Toronto data. First, consider the Tense/Morphology factor group. Contrary to the hypothesis of diffusion based on Labov (2007), Toronto and Thunder Bay are close to identical. In the middle age group, *be like* is favoured by the present tense in general. This is the same as Toronto. Assuming that it was these middle age speakers who first used *be like* in the 1980s, and assuming that the source of *be like* was Toronto, the Thunder Bay speakers have replicated this constraint. Furthermore, in the second generation, we see the same trajectory for *be like*. Though it begins in the general present tense, as in Toronto it has

---

30 Considering just *Present* and *Historical Present* in the two communities, there is a contrast in Tables 11 and 12. In Toronto, the factors have weights of .78 and .76 respectively. In Thunder Bay the weights are .56 and .71. I performed a Goldvarb run as if the factor *Past* was a “Knockout” to test the significance of the difference between *Present* and *HP* in the middle age groups. In Toronto, the factor group did come out as significance but with a very small range (*Present*.52; *HP*.49; Range 3). In Thunder Bay, the effect was also significant, but in the opposite direction (*Present*.43; *HP*.59; Range 16).
specialized to the historical present. Every instance of historical present forms in the younger generation co-occur with *be like*. Further, the past and present tenses do not constrain the realization of *be like* in a significantly different way from one another, suggesting that as in Toronto, *be like* is specialized to the HP.

A second consistency with Toronto is the strong Sex effect. Females in both speech communities in both generations favour *be like* and this effect strengthens through time.

However, the Thunder Bay quotative system is not a replica of the Toronto system. First, the 1st/3rd person effect is weak in the middle age group in Toronto is not found for Thunder Bay. The effect is not-significant. The effect is significant in the younger generation in Thunder Bay though with a range of only two.

More importantly, the Content of the Quote constraint significantly differs in the two communities. In the middle generation, the Content of the Quote constraint is not observed to significantly constrain *be like*. This can be interpreted as the imperfect replication of the complex linguistic structures associated with diffusion in Labov’s (2007) model - the constraint was not acquired. In the younger generation, the factor group is observed to significantly constrain the realization of *be like*. However, the effect is the reverse. Direct Speech favours *be like* (.52 factor weight) while Inner Dialogue disfavours *be like* (.33 factor weight).[^31]

Sali Tagliamonte (pc.) and Gerard van Herk (pc.) both point out concern regarding the small Ns in this factor group. Acknowledging this, I am cautious in interpreting the results as definitive observations of diffusion of change. However, Table 10 suggests that the types of narrative in Toronto and Thunder Bay may differ. Tagliamonte & D’Arcy (2007b:211) observe a recent increase in speakers “telling stories by regaling their audience to a running stream of their own inner thought processes.” However, only the oldest speakers in Toronto have rates of Inner Dialogue comparable to the Thunder Bay average (nine percent). Again, the scatterplot view of individuals in Thunder Bay is

[^31]: To the best of my knowledge, this pattern has never been reported for *be like*. Ferrara & Bell (1995, as reported in Tagliamonte & Hudson 1999), Tagliamonte & Hudson’s (1999) British and Ottawa data, and Tagliamonte & D’Arcy (2007b) all report that *be like* is favoured in contexts of Inner Dialogue. Ferrara & Bell (1999) even suggest that *be like*’s initial stage, whatever the variety of English, will favour Inner Dialogue and then expand to Direct Speech. However, in cases of linguistic diffusion, the only prediction about how an innovation will act in a new speech community is that the feature will be observably different in the new community because adults do not have the same language learning abilities as children (Labov 2007:371).
a more revealing picture. Figure 14 plots individuals by the proportion of their tokens that express Inner Dialogue. The apparent time Toronto data from Tagliamonte & D’Arcy (2007b:211, Fig. 4) is also plotted for comparison.\(^{32}\)

Figure 14
Scatterplot of individuals in Thunder Bay by proportion of Inner Dialogue tokens and Toronto mean

![Scatterplot](image)

Figure 14 shows that unlike the monotonic increase of Inner Dialogue observed in Toronto, Thunder Bay exhibits a different profile. The trendline suggests a decrease in Inner Dialogue in the community through time. Note that there are eight speakers in Thunder Bay who hover around the Toronto average but only three exceed that average. This perspective adds further evidence to the suggestion that Thunder Bay and Toronto have different narrative styles. Where Toronto is characterized by “the rising stylistic option of inner monologue in narratives of personal experience” (Tagliamonte & D’Arcy 2007b:211), the same cannot be said for Thunder Bay.

Tagliamonte & D’Arcy (2007b:211) argue that it was this rise in Inner Dialogue generally that lead to the rise of \textit{be like}. Tagliamonte & D’Arcy (2007b:211):

Thus, \textit{be like} filled a niche that already existed [increased Inner Dialogue], only later specializing into a device to grandstand the narrator as a

\(^{32}\) Speakers in Thunder Bay with less than 5 tokens were excluded here (N=3).
participant in his or her own stories. From there, the path was laid for it to develop into a narrative present marker in its own right.

There is no evidence for this developmental path in Thunder Bay.

In spite of the issue of small Ns in the Inner Dialogue cell,\(^{33}\) the finding that speakers in Thunder Bay differ from Toronto in the nuances of their story-telling is suggestive of an interpretation of diffusion under Labov’s model. Be like arrived in Toronto to fill a specific niche but arrived in Thunder Bay without filling such a niche. Whether by virtue of or epiphenomenal to the difference in narrative style in the two communities, the association with Inner Dialogue was not acquired by speakers in Thunder Bay. In this respect, the innovation displays a profile of imperfect replication of the abstract features underlying variation in the system, fitting the model of diffusion.

The Thunder Bay results are an interesting addition to the existing literature on be like. Though the Tense/Morphology factor group previously identified as constraining be like in Toronto is active in Thunder Bay, the Content of the Quote constraint, and the Person constraint (though to a lesser extent) differ.

The apparent time evidence presented in Figure 9 above indicates that be like was diffused to Thunder Bay because no one in the community over the age of 40 uses the variant. Furthermore, the demographic evidence suggests that Toronto is Thunder Bay’s source for be like. Following Labov’s (2007) model of diffusion, we would expect the distribution of be like observed in Toronto to be imperfectly replicated in the Thunder Bay data. To a point, this is what is observed. However, some pieces do not seem to make full sense within this model.

How can the seemingly perfect adoption of the complex tense constraint be explained? I can suggest three potential explanations: 1) diffusion acts differently above the level of phonology, 2) transmission and diffusion are not the only two types of

\(^{33}\) One wonders what effect more tokens would have. Given the present data set for Thunder Bay, we would need 12 more Inner Dialogue tokens of be like for this factor effect to be reversed (to what is observed in Toronto). Recall that Direct Speech favours be like with a proportion of 127/151 (=84.1 percent). Currently there are 10 token of be like in the Inner Dialogue cell and 4 tokens not of be like [10/14 = 73 percent]. Adding 10 tokens of be like would result in a proportion of 22/26 = 84.6 percent. Adding anything less would preserve the effect (e.g. 20/24 = 83 percent). Ten tokens seems like only a handful, but considering that there are only 10 tokens to begin with, we would need to double the size of the cell.
linguistic change and 3) changes due to diffusion are characterized by imperfect replication of only *non-universal* constraints.

The first explanation is that Labov’s model cannot be applied above the level of phonology. However, this is less than desirable. Variationist studies have consistently found that phonological and grammatical changes behave in the same manner. For example, Sankoff & Thibault (1981) observe that linguistic change at any level can be analyzed under the variationist framework, while Tagliamonte & D’Arcy (2009) found that the *incrementation model*, which was articulated using phonological variation, is readily applicable to changes above the level of phonology. For Labov’s (2007) characterization of diffusion to be applicable only to phonological changes would seriously limit its generality. We need another explanation.

Perhaps *be like*’s usage in Thunder Bay is not a result of the diffusion process or the transmission process at all, but rather some other type of change not addressed in Labov (2007). Labov (1994:84) discusses *communal change* in which “all members of the community alter their frequencies together or acquire new forms simultaneously.” Tagliamonte & D’Arcy (2007b:213) observe through a real-time analysis that the Toronto speech community has experienced communal change. Those speakers who are now middle-aged use *be like* more than 30 percent of time. *Be like* was only observed to occur 13 percent of the time in the same set of speakers in the mid-1990s (at that time in their teens). An adult shift in frequency is characteristic of communal change. However, such an interpretation is not consistent with the Thunder Bay data. In Thunder Bay the frequencies are similar to Toronto (see Figures 8 and 9) but the constraints are not exactly the same. Another type of change is *divergence*, which has its theoretical roots in Speech Accommodation Theory (Giles 1973). This type of change is characterized by speakers diverging “from a group they don’t like/approve of” (Nagy 2009). Examples in Welsh, Flemish and Fijian are reported in Niedzielski & Giles (1996:336). The behaviour of speakers in Thunder Bay with respect to *be like* does not fit this type of change either. Other than the observed female lead, associated with changes in progress (Labov 1972b), there is little evidence that *be like* has any social correlates in Thunder Bay, let alone it having a negative association with another community (Toronto) that would trigger differences in the quotative systems of the two communities.
The last potential explanation I will discuss holds that Labov’s model is applicable above the level of phonology but some reworking is required. Consider again the quotation from Tagliamonte (2006:246) reproduced below.

If the direction of effect of a factor group is shared by varieties, and the effect is not universal, this can be evidence that the varieties have inherited that constraint from a common source. On the other hand, where there are dissimilarities, this can be grounds for concluding that the phenomena in question belong to different linguistic systems. (emphasis added)

The universal nature of constraints is a crucial point and one that to the best of my knowledge has yet to be addressed. What are universal constraints on variation? How do they differ from community specific effects? How can the two be disentangled? The results of this paper may reveal a first approximation of an answer to these questions. The quotative system in Thunder Bay is characterized by sharing the Tense/Morphology constraint with Toronto, but it shows a difference with respect to the Grammatical Person of the Subject constraint and particularly the meta-pragmatic Content of the Quote constraint. If we take diffusion to be characterized by the imperfect replication of non-universal constraints, then perhaps the Content of the Quote constraint should be considered community-specific while the other constraints are universally connected with be like. However, Dion & Poplack’s (2005) examination of English speakers in Quebec City and Montreal, minority speech communities in a French dominant region, found that the variable grammars of 18-35 year olds in Quebec were nearly identical to the grammars of Tagliamonte & D’Arcy’s (2007) speakers of the same age group in Toronto. Dion & Poplack (2005) observe that all the constraints on be like are significant in the minority community and the hierarchy of effects is identical in all the factor groups except the morphology/tense group where unambiguous present tense with present morphology slightly favours be like (.56) where the same factor disfavours be like in Toronto (.36). Otherwise, the communities are identical, including the direction of effect of the Content of the Quote factor group. Both communities favour be like in Inner Dialogue. Why would one community where be like was diffused to (Quebec) be identical to Toronto while another community where be like was also diffused differ from Toronto? Also relevant to Dion & Poplack (2005) are the recent observations out of
London, England. Cheshire & Fox (2009) find the same effect for Content of the Quote reported here: Direct Speech favours *be like*.

This is the point of departure for this paper. Any definitive answer to such questions would require an examination of a number of different innovations in the progress of diffusion in a number of different varieties of English. This paper leaves these potential explanations open for future refutation or confirmation.
Section 7. Conclusions

This paper considered two morpho-syntactic changes in Canadian English. The nature of these changes was explored by applying Labov’s (2007) distinction between transmission and diffusion of linguistic change. I considered the community of Thunder Bay as a testing ground for this distinction. The rise of have to as a marker of deontic modality was argued to be a process shared by Thunder Bay and its most influential speech community, Toronto. Historical evidence suggests that the change from must to have to was in progress at the onset of Canadian English, and this change in progress was transplanted to Thunder Bay at the time of settlement. It has since been transmitted uninterrupted in the two speech communities and, as a result, identical systems are active in both. Consequently, transmission of morphosyntactic change fits Labov’s model of transmission, characterized by the “maintenance of complex language structures” (Labov 2007).

The nature of innovative be like in Thunder Bay is a less clear cut case. Labov (2007:349) characterizes diffusion as resulting in the “imperfect replication of abstract features of language structure” and one of the constraints on be like, the Content of the Quote constraint, is ranked differently in Thunder Bay than in Toronto. However, another constraint was replicated faithfully. These constraints suggest that the application of Labov’s transmission and diffusion model to changes above the level of phonology may not be a straightforward pursuit. This paper leaves open the question of why some constraints and not others are acquired through diffusion. However, we may be able to operationalize change caused by diffusion to disentangle universal constraints on variants from those that are community specific.

Leaving open a number of issues is not entirely satisfying. What can we say definitively? One observation is that (at least the deontic modals) in the community of Thunder Bay fits the consistent observations of homogeneity - a result of vast westward migration of Ontarians (Chambers 1991:91). Thunder Bay’s system is identical to the system observed in Toronto. The second point is that Northern Ontario is not immune to the world-wide spread of be like. Such newer changes might lead to the eventual diversification of Canadian English, as Chambers (1991:92) predicts. Certainly,
examining lesser-studied communities throughout Canada (especially beyond Ontario) is the future of research on Canadian English. Canadian English’s homogeneity and autonomy makes for the perfect laboratory for studying issues of language change. Extensive study of speech communities throughout Canada should shed light on the open-ended questions brought up in this forum paper.
References


