A Comparative Variationist Approach to Morphosyntactic Variation in Hexagonal and Quebec French

Laura Kastronic

A thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the
Doctorate in Philosophy degree in Linguistics

Department of Linguistics
Faculty of Arts
University of Ottawa

ABSTRACT

French spoken in Quebec (QF) is largely considered to be vastly different from Hexagonal French (HF), spoken in France, but this assumption has not been tested empirically. Systematic and accountable analyses of morphosyntactic features in HF are generally absent from the current body of linguistic research on French, which precludes the possibility of reliable comparison.

This dissertation addresses this gap through the systematic analysis of three different linguistic variables in spoken HF (the subjunctive mood, the expression of necessity, and the expression of pronominal indefiniteness), using the comparative variationist sociolinguistics framework, which seeks to account for variability in language by demonstrating that choices made by speakers are not random, but rather are constrained by various social and linguistic factors.

The analysis tests whether the strong community trends that deviate from the prescribed standard, shown to be operative in QF (e.g., Poplack, Lealess & Dion 2013), are also present in the source variety (HF) and if so, whether the underlying conditioning of the variability is parallel across the two varieties.

The results of this research reveal nearly identical underlying systems in both varieties in the case of the subjunctive mood and the expression of necessity. The most notable difference is the circumstantial proportion of specific lexical items (namely *falloir que*, which is strongly entrenched with the subjunctive mood in QF and is also the main variant used to express necessity), which is eventually linked to the different strategies employed in each variety to express indefinite necessity and indefiniteness more generally. The results of this research

highlight the importance of considering the consequences of one change on another and how variability in one area of the grammar contributes to the larger picture. It also represents the first comparative variationist analysis of morphosyntactic variation in QF and HF, and thus contributes to clarifying our understanding of the (dis)similarities between the two varieties, in addition to offering insight into the relationship between transplanted dialects and their source variety more generally.

ACKNOWLEDGEMENTS

Without the support of wonderful people in my life, this rewarding, yet challenging, journey may not have come to completion. I am truly grateful to all of them.

First and foremost, I would like to thank my supervisor, Shana Poplack, for her guidance, generosity, and valuable insight throughout this entire process. It has been an honour to be trained and mentored by a world-renowned and highly decorated sociolinguist, Canada Research Chair, and pioneer in our field. The world-class training I received from her not only allowed for the completion of this project, amongst several others, but it will continue to serve as the foundation for my future work. I am also grateful for having been granted access to her vast collection of one-of-a-kind speech corpora, particularly the two used in the context of this study: *Corpus du français parlé à Ottawa-Hull* and *Le français en contexte: milieux scolaire et social.*

I would also like to extend gratitude to those who compiled the Hexagonal French corpora which form the basis of this study. For the *Corpus de Français Parlé Parisien des années 2000,* this includes Sonia Branca-Rosoff, Florence Lefeuvre, Serge Fleury, and Mat Pires. For *Enquêtes Sociolinguistiques à Orléans,* I would like to thank the project coordinators, Olivier Baude and Céline Dugua, and the members of Laboratoire ligérien de linguistique at l'Université d'Orléans who, recognizing the importance of spoken corpora to linguistic research, took the original data collected by British academics in the 1960s and expanded upon it more than forty years later. I would like to thank both Céline Dugua and Olivier Baude for their help in granting me access to the data, in particular M. Baude, who was extremely patient with me and graciously addressed all of my inquiries without hesitation. Most importantly, I would like to say how grateful I am that both of these corpora are available for consultation online. Without

access to such high-quality data in both France and Quebec, this project would not have been possible.

I sincerely appreciate the time and effort that each of my thesis evaluation committee members dedicated to providing me with their feedback on this study: Éric Mathieu, Ana Arregui, Stephen Levey, and Julie Auger. Each offered an array of insightful and thought-provoking comments, many of which I have addressed in this final version.

I am also indebted to Nathalie Dion, who has been a great support to me since embarking on this journey. No matter how busy, she always had time to discuss results, ideas, and offer her feedback. Past and present labites and other colleagues have contributed to a rewarding, supportive, and fun research environment: Kim, Nahed, Salvio, Basile, Rick, Yukiko, Mélissa, and Suzanne, amongst others.

The Linguistics Department at the University of Ottawa has been a wonderful place to complete my doctoral degree. I would like to thank all of the faculty members for contributing to such a fruitful learning environment, in particular Éric, Laura, Tania, and Stephen, all of whom were available for a chat at a moment's notice.

I have also been extremely lucky to be surrounded by wonderful people in my personal life. I want to thank my mom, my dad, and my brother for their endless love, encouragement, and patience. I am also grateful to Cecille for her continuous support and for always cheering me on. My close friends, Ashley and Dunni, were both there for me in countless ways throughout this entire process, which included listening to me ramble on about my research and helping me get my mind off everything with a night out (well, many nights out!) and lots of wine. I can't forget to mention our kitties, Charlie and Yoda, who both passed away while I was writing this. They

forced me to look up from my work (by lying on top of it!) and always made me smile. Thankfully, our new cat, Ludo, has taken over the job.

Lastly, but most importantly, I would like to thank my husband and life partner, Greg, whose support of me has never wavered. He has stood by my side at each step of the way, through all of the ups and downs, selflessly, patiently, and loyally. He fills my life with love, laughter, and true friendship, for which I am infinitely grateful.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

CHAPTER 2: METHODOLOLGY

CHAPTER 3: THE SUBJUNCTIVE

CHAPTER 4: THE EXPRESSION OF NECESSITY

Chapter 1 INTRODUCTION

French spoken in Quebec (QF)[1] is widely accused of being inferior to Hexagonal French (HF). Its many characteristic (often stigmatized) features are believed to have originated and developed on Canadian soil, due either to separation from the metropolis, or to contact with English (Cajolet-Laganière & Martel 1995). While it is widely assumed that the two varieties are substantially different, this belief has not been tested empirically in a methodologically sound manner. This gap will be addressed in this dissertation through the creation of a benchmark of contemporary HF to which contemporary QF can be compared in an effort to test, using the comparative variationist method (Poplack & Tagliamonte 2001), whether the strong community trends operative in French spoken in Quebec, as elucidated by the research of Poplack (1990; 1992; 1997; 2015; 2009 etc.) and many others (Miller & Dion 2009; Poplack & Dion 2009; Poplack & St-Amand 2009; Elsig & Poplack 2006; Thibault 1991a; Thibault 1991b; Sankoff & Vincent 1977; Blondeau 2008; Laberge 1977 etc.) are also present in the source variety and whether they are subject to the same constraints on variability. The variables under analysis are the expression of the subjunctive, the expression of necessity, and the expression of pronominal indefiniteness.

1.1 French in Quebec

Starting in the 16th century, European expansion across several different continents resulted in the spread of many European languages to the colonies, including Spanish, Portuguese, French, and English. This period of European colonialism and subsequent language

[1] Some scholars refer to French spoken in Quebec as Laurentian French, a term which encompasses not only QF but also all of the varieties which originated along the St. Lawrence River and spread to various regions throughout Canada and the United States (Côté 2012). Throughout this study, the focus is specifically on French spoken in the Hull region of Quebec, and thus I refer to it as Quebec French (QF).

transplantation resulted in the creation of new dialects of the European languages (e.g., Brazilian Portuguese, Mexican Spanish, Australian English, and Canadian French) which are often said to differ from their original source (Trudgill 2008). This research focuses solely on the French Canadian case, specifically French spoken in the Ottawa-Hull region (Figure 1.1), along the Quebec/Ontario border.

Figure 1.1: Map of Ottawa-Hull region in Canada (Google 2016).

1.1.1 History of Quebec French

French spoken in Quebec is considered to have its roots in the regional dialects of Hexagonal French, and was putatively "modified by a long period of intense contact with English and by adaptation to North American conditions" (Ager 1990:102). In the early 17th century, French colonists settled two areas in North America: Acadia (along the Atlantic coast, now comprised of the Atlantic provinces: New Brunswick, Prince Edward Island, and Nova Scotia) and *Nouvelle-France* 'New France' (along the St. Lawrence River, now the province of Quebec) (Blanc 1993:240). French colonists who settled along the St. Lawrence River originated from all over France (Charbonneau 1990; Corbeil 2007), but the majority came from Western

France, where regional dialects were abundant (Ager 1990:102; Charbonneau 1990). According to Table 1.1, approximately 9000 immigrants settled in *Nouvelle-France* between 1608 and 1765, many of whom spoke French, regional dialects, or a combination of both (Corbeil 2007).

	Période d'arrivée				
Région d'origine	1608-1679	1680-1699	1700-1729	1730-1765	Ensemble
	Nombres absolus				
Bretagne	95	72	84	225	476
Normandie	853	107	56	339	1 355
Paris	665	134	174	350	1 323
Loire	319	68	77	181	645
Nord	97	18	28	172	315
Est	127	35	69	480	711
Ouest	1 107	339	285	386	2 117
Centre	78	68	45	136	327
Sud	130	156	157	714	1 157
Indéterminée	292	117	119	12	540
Total	3 763	1 114	1 094	2 995	8 966
	Nombres relatifs				
Bretagne	2,5	6,5	7,7	7,5	5,3
Normandie	22,7	9,6	5,1	11,3	15,1
Paris	17,7	12,0	15,9	11,7	14,8
Loire	8,5	6,1	7,0	6,1	7,2
Nord	2,6	1,6	2,6	5,8	3,5
Est	3,4	3,2	6,3	16,0	7,9
Ouest	29,4	30,4	26,1	12,9	23,6
Centre	2,1	6,1	4,1	4,5	3,7
Sud	3,4	14,0	14,3	23,8	12,9
Indéterminée	7,7	10,5	10,9	0,4	6,0
Total	100,0	100,0	100,0	100,0	100,0

Table 1.1: Origins of settlers in New France
(*nombres absolus* = raw numbers, *nombres relatifs* = percentages) (Charbonneau 1990:53).

Despite their diverse origins, the mixture of French and regional dialect speakers did not settle into separate linguistics communities, which resulted in the emergence of a common language "not unlike Standard French", according to some (e.g., Blanc 1993:241). French was the dominant form of expression in *Nouvelle-France* until the British took over in the late 18th century through their defeat of the French on the Plains of Abraham (Corbeil 2007:63), at which point *Nouvelle-France* (renamed *Quebec*) became a British colony, which subsequently led to

the infiltration of English into many aspects of society for many decades (Blanc 1993). By 1960, English was the dominant language in business, and immigrants to the province were choosing to learn English instead of French in hopes of attaining a higher socioeconomic status (Corbeil 2007; Dickinson & Young 2003).

The favourable status of English in Quebec by the second half of the 20th century led to the implementation of a number of language laws designed to reinstate and secure the status of French by increasing its presence in various domains (Ager 1990:101–102; Corbeil 2007), many of which remain in place today. Not only has the status of QF been the subject of controversy, so has its quality. In fact, French spoken in Quebec has long been subject to normative influence, inspired by its use in the source variety (HF) (Blanc 1993).

1.1.2 Reactions to Quebec French

The supposed inferiority of QF in comparison to HF is largely due to its many purportedly distinct (and frequently stigmatized) features. In order to access the general reaction towards QF, Cajolet-Laganière and Martel (1995) created a corpus that was comprised of language-related comments in *La Presse* (a popular Quebec newspaper) between 1960 and 1993 and found that the overall evaluation of QF was overwhelmingly negative. Interestingly, language commentary increased twofold after 1980, which coincides with the implementation of several language laws designed to preserve the French language in Quebec. Nearly all (90%) of the commenters viewed QF in a negative light and these impressions were not unique to regular readers (i.e., laypeople). In fact, about half of comments originated from teachers, linguists, and other ‘language specialists’. The general reaction towards French in Quebec was that it was a degraded and vulgar dialect, as is summed up by the following quotes:

« *Notre français parlé, c'est, sauf exception, une souffrance pour toute oreille exercée à la correction, à l'harmonie. C'est le plus souvent une langue* ***pauvre, dégradée*** *par un nombre considérable de déformations*. » (Le 25 janvier, 1968, T.B., La Presse 1968:4).

'Our spoken French is, without exception, unbearable to the well-trained ear which is attuned to correction and harmony. It is usually a poor language, degraded by many considerable deformities.'

« *Notre misère linguistique est profonde, les Canadiens français sont* ***vulgaires***. » (Le 22 avril, 1960, Jean-Marie Laurence, La Presse 1960:9).

'Our linguistic misery is deep, French Canadians are vulgar.'

The negative commentary towards Quebec French was not restricted to the general population, but was also directed at lawyers, teachers, doctors, politicians, and other well-respected figures in society (ibid., 24).

« *Quand on entend des hommes d'affaires, des médecins, des notaires, des avocats, des ingénieurs et même des prêtres et des universitaires s'exprimer comme des charretiers, eh bien, que voulez-vous, on ne peut pas donner tort au charretiers.* » (Le 5 mai, 1960, Candidus, *La* Presse 1960:4)

'When we hear business men, doctors, notaries, lawyer, engineers, and even priests and academics speak like a carter (fig. 'speak like a sailor'), well, what can you expect, we can't blame the carters.'

By the end of the 20th century in Quebec, the general opinion of QF was such that it was a corrupted version of a once pure and correct language (HF), and this was largely attributed to contact with English (Cajolet-Laganière & Martel 1995).

Despite such reactions to QF, which many would argue persist in varying degrees even today, little comparative quantitative work has been done to confirm or refute assumptions about its supposed degraded nature. This is one of the motivations for the present work.

1.2 Transplanted dialects

Transplanted dialects, like QF, are said to differ from the source variety in many ways. What follows is a brief discussion of how they are formed, some of their general linguistic characteristics, as well how they can be modelled theoretically.

1.2.1 The tree model

Schleicher's (1861) tree model of linguistic comparison (*Stammbaum* model) has been crucial in comparative and historical linguistics in mapping the familial relationships between languages across time (Fox 1995). One clear advantage of this model is that it allows for representation of "relative closeness" of two languages and can easily represent multiple dialects (or dialect split) by adding "finer and finer branches" (Anttila 1989:302).

However, the tree model has been criticized practically since its inception due to its inability to account for outside influence. As is suggested by Schuchardt (1868): "each word has its own history" (cited in Campbell 2013:188) and thus the linguistic elements which make up a language or variety can be "the result of various influences from various directions" (ibid., 189). Since this model only allows a language or dialect to have a single parent, it falsely predicts that it cannot be influenced by any other branch of the tree (Fox 1995). In other words, it is unable to account for language change due to contact with other languages or dialects (Fox 1995) and linguistic diffusion (Anttila 1989). As Wolfram and Schilling-Estes (2003:721) suggest, it is "a highly idealized representation of a far messier reality."

Using an example from Modern English, Figure 1.2 (reproduced from Fox 1995) illustrates the inadequacies of the tree model in representing the relationship between a language

variety and its source. In (b), the prototypical mapping of the origins of Modern English is shown, while in (a), a more realistic representation is presented.

Despite its limitations, the tree model easily conceptualizes the relationship between HF and QF in terms of the origins of the latter. But the question remains about what happened next, after the split. This is a question for which the tree model provides no answers.

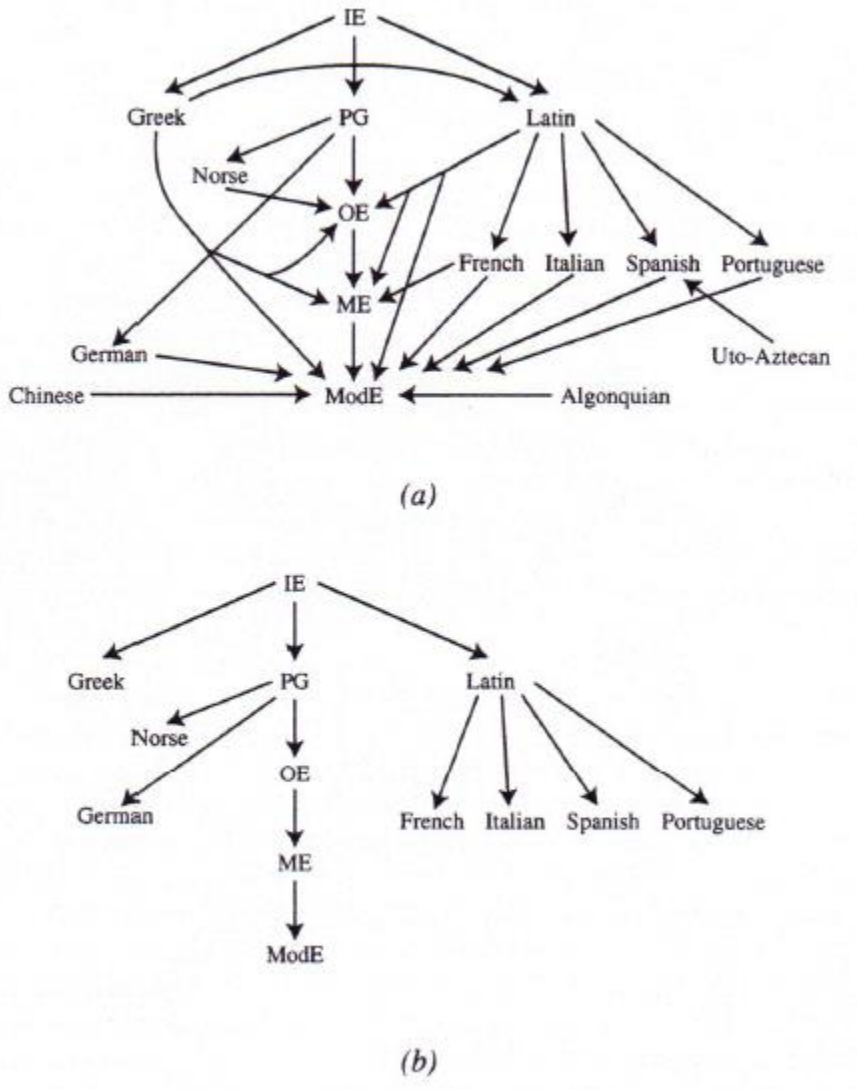


Figure 1.2: Origins of Modern English (adapted from Fox 1995:124).

The shortcomings of the tree model with regards to linguistic change, first identified in the late 19th century, inspired the development of wave theory (Schmidt 1872), which suggests that linguistic change spreads outwardly from a central point to surrounding areas and thus can account for dialect- and contact-induced change (Campbell 2013:188–189). The gravity model (Trudgill 1974) later replaced wave theory in terms of its capacity to explain the diffusion of linguistic change by incorporating both population size and geographic distance. However, many linguistic phenomena remained unexplained by these models, such as parallel innovations which occur in geographically distinct varieties but not in the source, as noted by Trudgill (2004) for southern hemisphere Englishes.

1.2.2 Dialect contact and dialect mixing

There is consensus that dialect contact occurred in all transplanted European varieties due to the fact that European settlers rarely originated from a single location in the homeland (Trudgill 2008). Not all scholars agree that dialect contact necessarily led to dialect mixing, a concept which presupposes mixture of dialects from the source variety as the basis for the newly formed transplanted varieties (Trudgill 2008). Some historical linguists insist that certain transplanted varieties have monogenetic origins (i.e., they originated from a single source) and, therefore, did not undergo the process of new dialect formation (Trudgill 2008:241–242), as is the case for Rivard (1914) who suggests that the origins of QF can be found in the dialects of Normandy (cited in Trudgill 2004:7).

In terms of the origins of QF, some believe that settlers mostly spoke regional dialects (e.g., Picard, Franco-Provençal) and had little to no knowledge of French (e.g., Barbaud 1994; Dulong 1973), while others (e.g., Chaudenson 1994; Mougeon & Beniak 1994; Poirier 1994) assert that most spoke at least some degree of French. The latter explanation is supported by the

fact that *les habitants* (settlers of New France) originated from various, mostly urban (Charbonneau & Guillemette 1994), locations in North-West France (Table 1.1) and had a literacy rate much higher than that of the general population (Trudel 1973, cited in Barbaud 1994). In addition, Morin's (1994) study of the source of pronunciation in QF clearly shows that many present-day pronunciation patterns are based on either internal development or the Parisian norm at the time of colonization.

Despite the fact that there is no agreement on which regional dialects contributed to the creation of contemporary QF, the importance of dialect contact is generally agreed upon, as expressed by Mougeon and Beniak (1994:25):

> "The particular conditions associated with emigration to New France (Canada) and with life in the new colony on the St. Lawrence led to a unique mixture of central French, varieties of regional French and perhaps even of certain *patois*…From this mixture there developed a new form of the French language…a unified, coherent, distinctive Quebec variety." (translation from Trudgill 2006:269)

While the question of the exact origins of QF is interesting, it is beyond the scope of this study[2]. The purpose of this study is to determine the degree to which French spoken in Quebec today actually differs from the contemporary source variety, if at all.

1.2.3 Characteristics of transplanted dialects

Transplanted dialects are often said to diverge from the source variety at many levels of linguistic structure. This can include lexical, phonetic, morphological, and syntactic differences. Britain (2004:42–43) outlines three characteristics of dialects formed by means of "relocation diffusion" and dialect contact, including the presence of unmarked majority variants resulting from levelling, a "structurally less complex" linguistic system, and "the emergence of

[2] For a detailed summary of several of the theories regarding QF's origins, see Mougeon & Beniak (1994).

interdialect forms" which did not exist in the source variety. Levelling refers to the replacement of irregular variants with regular ones and those which are most likely to emerge as the majority variant tend to be socially neutral and the most socially and geographically diffused (Britain 2004:42; Schreier 2003:12). Decomplexification (also referred to as simplification) occurs when "a given property of variety X, no matter if phonetic/phonological, lexical or grammatical, is subject to less variation than it originally was in the input varieties" (Schreier 2003:13). In other words, there is usually a reduction in the number of variants. Finally, the emergence of new forms may occur in the transplanted dialect which are absent from the source variety or varieties (i.e., interdialect forms) (Britain 2004:43).

Trudgill suggests differences between a dialect and its source can be attributed to one or more of the factors outlined below (cited directly from 2004:1–2):

Factor 1: The adaptation of the transplanted dialect to new realities which were not present in the source variety.

Factor 2: Linguistic changes which occur in the source variety but not in the transplanted dialect[3].

Factor 3: Linguistic changes which occur in the transplanted dialect but not in the source variety.

Factor 4: Contact of the transplanted dialect with indigenous languages.

Factor 5: Contact with other (European) languages.

Factor 6: Dialect contact within the transplanted variety.

In terms of HF and QF, many of the attested differences are lexical in nature and thus attributable to factors 1, 4, and 5. Such differences include not only borrowings from English (*anglicismes*), but also borrowings from aboriginal languages (e.g., *moccasin, tomahawk,*

[3] This scenario is often labelled in the literature as *colonial lag*, which Trudgill (2004:34) suggests refers to situations in which the transplanted dialect lags behind the source variety when it comes to linguistic change.

sachem) with which the earliest speakers of QF necessarily came into contact, as well as neologisms created to account for realities of life in North America that were different from France (e.g., types of fauna: *achigan, ouananiche, ouaouaron*) (Cajolet-Laganière & Martel 1995:49–50). Lexical differences between the two varieties also stem from QF's retention of certain archaic lexical items from 17th and 18th century HF (e.g., *astheure* 'right now'; *tanner* 'to badger'). In addition, QF is well known for its feminization of certain titles, which can be accomplished in several ways, including by adding a feminine pronoun (e.g., *une journaliste* 'a journalist'), use of the word *femme* (e.g., *femme-ingénieur* 'female engineer') to the otherwise masculine word, or even by creating a feminine form (e.g., *practicienne* 'medical practitioner') (de Villers 2008)

Aside from lexical differences, there are clear phonetic differences between the two varieties that have also been attested. For example, in some regions and/or varieties, QF retains the archaic pronunciation of *–oi*, as in *moi, toi, voir, etc.* (these were pronounced *moué, toué, voué* amongst the elite under the reign of Louis XIV and Louis XV) (Juneau 1972, cited in Cajolet-Laganière & Martel, 1995:47). Since this pronunciation is largely believed to be absent from HF speech[4], this would suggest that the source variety has changed, while the transplanted dialect retained an archaic feature, corresponding to Trudgill's third factor. Another example is the affrication of /t/ and /d/, which is a widespread feature of QF (Juneau 1976 cited in Cajolet-Laganière & Martel 1995:57), and appears to be the result of change in the transplanted dialect, thus corresponding to Trudgill's second factor.

While QF vocabulary and pronunciation have received considerable attention, much less is known about morphosyntactic characteristics. While some differences have been noted (e.g.,

[4] It is of course possible that this pronunciation is used in some variety of HF; however, it is commonly cited as a difference between (certain varieties of) QF and HF.

interrogatives ending in [ty], *dis-y* instead of *dis-lui*, *je vas* instead of *je vais*, etc.) (Mougeon 1998; Cajolet-Laganière & Martel 1995), most of these assertions are not based on quantitative results. In fact, many of these are also attested in HF, as is discussed below. Qualitative observations can be valid, but can also be tainted by the observer's own biases which can result in the (often incorrect) assumption that variation is change. Without quantitative analysis, it is impossible to determine to what degree a particular linguistic feature has actually penetrated the language.

The identification of certain linguistic features believed to be unique to QF has led some researchers to characterize it as a simplified dialect, a common result of dialect contact/mixing (e.g., Britain 2004). For example, Chantefort (1976:31) claims that the disappearance of the *tu/vous* ('you') opposition for politeness, the increased use of *on* instead of *nous* as a first person plural pronoun ('we'), the increased use of the periphrastic future at the expense of the inflected future (*je vais aller* vs. *j'irai* 'I will go'), and use of the indicative in place of the subjunctive (amongst others) all suggest that QF has undergone decomplexification in comparison with the source variety.

One of the problems with such assertions is that there is a clear lack of comparative quantitative studies of HF and QF, particularly in the area of morphosyntax, to support them. All of the above-mentioned morphosyntactic features listed by Chantefort as characteristic of the simplification of French spoken in Quebec have been the focus of quantitative analysis in QF, and their presence has been confirmed in most cases. For example, Laberge (1977:132) found that *nous* only represented 2% of definite first person plural expression and had almost entirely been replaced by *on.* Thibault (1991a) found that the use definite *tu* (in variation with *vous*) nearly doubled between 1971 and 1984 (31% > 59%), while Poplack & Dion (2009) found that

the use of the periphrastic future increased from 56% to 73% between 19th and 20th century QF. However, in each case, variation was conditioned by a multitude of factors, which represents an argument against the simplification hypothesis. In the case of the subjunctive, Poplack, Lealess & Dion (2013) found that its use had actually risen over time, contrary to the claim that it is being replaced by the indicative.

Another argument against the simplification hypothesis is the presence of similar patterns of variation and change in HF. In the rare cases where there is a clear point of comparison (i.e., quantitative analysis of the same variable using the same methodological approach), it is clear that variant distribution and its internal conditioning (where available) is quite similar between the two varieties. For example, Roberts (2012) found that the periphrastic future was the majority variant in HF (59%) despite it being used less frequently than in QF (73% in Poplack & Dion 2009). In addition, he found that variant choice in HF was conditioned by internal (linguistic) factors in a manner nearly identical to QF, which speaks to deep-rooted similarities between the two varieties. Even in cases where the methodology is not comparable, it is clear that many of the features said to be unique to QF are also present in the source variety, such as first person definite *on* in place of *nous.* Coveney (2000) reports a rate of only 4.4% *nous* in a regional variety of HF, which is not dissimilar from Laberge's (1977) finding of 2%.

As was mentioned above, there is a lack of comparative quantitative studies of HF and QF, but this is not to say that there are no studies that examine variability in contemporary HF. Most of them, however, do not offer an explicit point of comparison of the underlying grammars of morphosyntactic variables in QF and HF, either due to differing methodological approaches or because they focus on regional variation (e.g., Villeneuve & Auger 2013; Coveney 2000; 1996). Some analyze only phonological variables (e.g., most studies stemming from the PFC project-

Durand, Laks & Lyche (2002); Coveney (2001), etc.), many are purely descriptive (e.g., Togeby 1974; 1966; van der Molen 1923), and amongst those that are quantitative, most focus only on rates of occurrence or social factors conditioning the variability (e.g., Coveney 2000; Lagerqvist 2003) as opposed to the linguistic factors and their respective strength and ordering, which together, constitute the underlying grammar, and can only be revealed through use of the variationist method (outlined in Chapter 2). In addition, due to the fact that many studies of HF focus on regional varieties (e.g., Villeneuve & Auger 2013; Coveney 1996; Coveney 2000), they are not necessarily representative of mainstream, contemporary HF. This study will aim to fill this gap through a horizontal (i.e., synchronic) comparison of mainstream spoken QF with mainstream spoken HF.

1.2.4 Variationist studies of Quebec French

Much is already known about the internal structure of French spoken in various regions of Canada[5], due in part to the existence of large, synchronic corpora of informal, spontaneous speech. This is particularly true of French spoken in Quebec, which has inspired the creation of many such corpora. The first scientifically constituted one is the Montreal French corpus (Vincent, Laforest & Martel 1995; Thibault & Vincent 1990; Sankoff & Cedergren 1972), which consists of spontaneous speech collected at three different points in time (1971, 1984, 1995), and has inspired dozens of variationist studies of Montreal French on a range of linguistic variables, such as /r/ (Sankoff & Blondeau 2007), future temporal reference (Sankoff & Evans Wagner 2006), and various discourse markers (Thibault & Daveluy 1989), amongst many others.

Our knowledge of French spoken in the capital region of Canada, which is the focus of the present study, is also quite extensive, due to years of variationist research undertaken by Dr.

[5] For a fairly detailed overview until the turn of the century, see Thibault (2001).

Shana Poplack and her associates at the Sociolinguistics Laboratory at the University of Ottawa. Several morphosyntactic features of QF have been analyzed using three large corpora of spontaneous spoken French, totalling over 4 million words (Poplack & Bourdages 2008; Poplack & St-Amand 2009; Poplack 1989). These corpora offer an unparalleled window into QF over the span of nearly a century and a half (19^{th} > 21^{st} century) at various points on the social and stylistic continuum. Some of the features studied through the use of these corpora include preposition stranding (phrase-final positioning of prepositions) (Poplack, Zentz & Dion 2012), the subjunctive (Poplack 1990; Poplack, Lealess & Dion 2013), future temporal reference (Poplack & Dion 2009; Poplack & Turpin 1999), possessive prepositions (Miller & Dion 2009), amongst many others.

Recurring themes in the multitude of studies of morphosyntactic (and phonological) variation in QF are: 1) that variability in spoken language does not reflect the prescriptive norm, and 2) there are several intrinsic "hidden" constraints on the variability that cannot be intuited and have not been formally taught (i.e., the *underlying grammar*). Despite this extensive body of work on QF, little has been revealed about its perceived differences from the source variety. Are the patterns identified in these studies unique to QF? Has change taken place in QF, but not in HF, or vice-versa? Is it perhaps the case that the two varieties resemble each other more than what most assume? These are some of the questions that inspired this research.

1.2.5 Language contact or internal evolution?

Differences between HF and QF would suggest that the two varieties evolved independently of one another along different paths, while similarities would likely be the result of shared retentions or shared innovations, the latter resulting from either "parallel but independent developments" or diffusion from one variety to the other (Trudgill 2006:266; 2004).

Shared retentions represent the strongest evidence to suggest that two or more varieties were inherited from a common source (Brugmann 1884). In the case of shared innovation, differences would only be discernible in the rate of change across the two varieties (i.e., if change in one variety is progressing at a slower/quicker rate than the other). Trudgill suggests that parallel developments in varieties inherited from a common source exemplify Sapir's (1921:172) notion of *drift*, which is based on the idea that "languages long disconnected will pass through the same or strikingly similar phases". In other words, it is unsurprising that multiple varieties originating from a single source tend to evolve in similar ways. Sapir (1921:150) acknowledges that even without dialect split, a language will continue to "move away from any assignable norm" and continue to develop to the point that it will be unrecognizable when compared to its earliest stage.

Despite this, independent evolution of two geographically distinct varieties along different paths is commonly attributed to contact with other languages (an explanation often invoked with regards to QF), often without rigorous analysis to support such assertions (Poplack & Levey 2010). In theory, contact-induced change can arise in almost any part of the language and between any pair of languages. In fact, it has been suggested that (e.g., Thomason 2001) in the right circumstances, any level of linguistic structure is susceptible to change. In addition, contact-induced change need not involve only the incorporation of a new linguistic feature into the recipient language. It can also involve the increase or decrease of an existing feature as a result of contact with another language which exhibits either more or less of said feature (Silva-Corvalán 2008).

In light of the widespread and often unfounded claims regarding contact-induced change in general, and the supposed influence of English on French spoken in Quebec, Poplack & Levey

(2010) outline methodological steps that can be used to demonstrate whether or not change is actually due to contact. A first step involves ensuring that change of any sort has actually taken place and that variation is not being construed as change. After change has been established, it must be demonstrated that differences across two varieties are due to one variety's contact with another language, and not internal evolution. Steps involved in establishing the source of the change (i.e., contact vs. internal evolution) involve the comparison with an earlier pre-contact stage of the language, with a non-contact variety, and with the source language.

Each time an accountable methodological approach has been adopted, results have revealed that in cases where it is plausible, contact with English fails to account for language variation and change in QF. For example, multiple studies of the subjunctive mood (e.g., Poplack, Lealess & Dion 2013; St-Amand 2002; Poplack 1992; 1990) found that, although its use in QF is structured and regular, it is far removed from prescriptive and theoretical accounts. However, there was no evidence to suggest that these divergences from the 'standard' were the result of contact with English, as suggested by some (e.g., Laurier 1989). This was evidenced in part by the absence of any effect of bilingual proficiency (Poplack 1997). If contact with English were, in fact, influencing subjunctive use in QF, this should be most apparent amongst the most proficient bilinguals, which was not the case.

Another variant analyzed in QF was the often stigmatized stranded preposition. Its use in QF is often blamed on contact since it is a widely accepted construction in English (King 2000) and is prescriptively prohibited in French, but Poplack, Zentz & Dion (2012) showed that it actually developed via analogical extension from the "orphan preposition", a structure native to French.

Elsig & Poplack (2006) compared the interrogative system in 19th and 20th century QF to that of 17th century HF, which was construed as a window on the state of French that was brought to Canada by the original settlers. They found few changes in QF across time and even showed that the QF interrogative system was actually quite similar to that of HF at the time of colonization. This suggests that it was not QF that changed over time, as many believe, but rather HF.

In sum, the assumption that QF differs from HF due to contact with English is largely unfounded, as is the assertion that it is a simplified version of the source variety. Such claims, when tested at all, stem from a variety of methodological shortcomings, including qualitative observations and comparison with an idealized standard, neither of which can offer reliable insight into the supposed dissimilarities between these two varieties of French. The results of this research will provide further evidence to refute these claims by adding a horizontal point of comparison with the source variety (i.e., contemporary HF).

1.3 Purpose of this study

The assumption that QF is both different and inferior to HF is longstanding despite a lack of accountable comparative analyses to support such a claim, particularly in the area of morphosyntax. This study aims to fill this gap by investigating the use of three morphosyntactic linguistic features through a horizontal comparative variationist analysis. It is important to note that in the case of morphosyntactic differences enumerated above (e.g., higher level of [ty] interrogatives, periphrastic future, less subjunctive, etc.), most of the forms are often present in both varieties thus signalling a quantitative, not qualitative, difference which need not necessarily be indicative of change. While quantitative differences can also be indicative of change, it is important to note that rate discrepancies may appear due to any number of reasons,

many of which are socially motivated (e.g., stylistic register), which is why it is crucial to also consider the conditioning of the variability (i.e., the grammar underlying the choice of one form over another). Following the above-mentioned analyses of QF, the systematic comparison of the conditioning of the variability will represent a major component of this dissertation and will be described in great detail in the next chapter.

1.3.1 Research questions

Despite the extensive research that has painted a clear picture of the many characteristic linguistic features of French in Canada, and specifically QF, crucial information is still missing: Are these features also present in the source variety? If the same type of variability does exist across both varieties, is there continuity in the linguistic conditioning of this variability? These questions, which will form the basis of this study, have remained largely unanswered. The systematic analysis of contemporary HF is crucial to gaining a clear understanding of the characterization and development of QF since it is 1) the source variety, and 2) has not been subjected to the same level of contact with English as QF.

While there is no shortage of spoken corpora for multiple varieties of Canadian French[6], few exist for HF. Fortunately, some have been collected in recent years (e.g., Branca-Rosoff et al. 2012 for Parisian French; LLL-Orléans 1968-2013 for Orléans French; Durand, Laks, & Lyche 2002 (PFC) for various varieties of European French; etc.), but they generally have yet to be exploited systematically from a variationist perspective, particularly for the morphosyntactic variables that are of interest here. This dissertation research addresses this gap through 1) the assembly of a large repertoire of three pre-existing contemporary spoken HF corpora; and 2) the

[6] For example: Poplack & Bourdages (2008) and (Poplack (1989) and for French in the capital region of Canada; Vincent, Laforest & Martel (1995), Thibault & Vincent (1990) and Sankoff & Cedergren (1972) for Montreal French; Mougeon & Beniak (1991) for Ontario French; Durand, Laks, & Lyche (2002) (PFC) for various varieties of French spoken in Canada, etc.

systematic analysis of specific morphosyntactic variables within these corpora, making use of the comparative variationist framework (Poplack & Tagliamonte 2001).

A horizontal comparison between contemporary QF and HF will shed light on current morphosyntactic (dis)similarities across the two varieties, which will, in turn, lead to a better understanding of the trajectory of development of both QF and HF. If the two varieties exhibit the same patterns of variability and associated constraints, one can assume that they were inherited from a common source and, therefore, both exhibit the shared retention or shared innovation of certain linguistic features. If they do not resemble one another, this would suggest that at least one variety must have changed. Could it be the case that QF actually retained features of the source variety, while HF changed over time? Or did they both diverge independently?

In this dissertation, three areas of the grammar (the expression of the subjunctive, necessity, and pronominal indefiniteness) are investigated in both QF and HF, synchronically (i.e., horizontally) and diachronically, using a systematic and accountable methodological approach (outlined in Chapter 2). In Chapters 3 and 4, a detailed analysis of these three linguistic variables is presented, including how the results from one the inquiry into the subjunctive led to the analysis of the expression of necessity, which, in turn, led to the study of pronominal indefiniteness. In Chapter 5, the findings of this study are summarized and discussed in the context of the bigger picture.

As will be demonstrated over the progression of this dissertation, there are remarkable similarities between the two varieties of French with respect to the morphosyntactic features under investigation here. Moreover, it is clear that both varieties are undergoing the same trajectory of change in more than one area of the grammar, albeit at different rates. These

findings refute claims regarding the influence of English on QF as well as those that characterize QF as a simplified version of the source variety.

Chapter 2 METHODOLOGY

2.1 Variation Theory

This research will be carried out within the comparative variationist sociolinguistics framework (Poplack & Tagliamonte 2001). As a combination of both the variationist and comparative methods, this framework is ideal for any cross-variety linguistic analysis because it allows the analyst to see past (dis)similarities between two cohorts by going beyond simple rates of occurrence in order to identify and then compare the underlying behaviour of any given linguistic feature.

The variationist method seeks to account for language variation and change through the systematic and accountable analysis of (spoken) language. A key construct of the variationist framework is the *linguistic variable* (Labov 1982; 1972; 1966), which "comprises a set of variants among which speakers alternate in expressing a given meaning or function" (Poplack & Levey 2010:398). In other words, speakers chose between alternate ways of expressing the same thing. These alternations only occur in specific contexts, which must be precisely identified in order to accurately account for variability. The process of identifying these specific contexts is referred to as *circumscribing the variable context* and is central to any variationist study. In circumscribing the variable context, one must adhere to the *Principle of Accountability*, which emphasizes the importance of considering not only the contexts in which a particular variant *does* appear, but also those contexts in which it *does not* appear, but could have (Labov 1972:72). The combination of all of these contexts (those in which it does appear as well as those

in which it could have appeared) represents the *variable context* and each individual item within the variable context is called a *token*.

Structured heterogeneity (or "orderly heterogeneity" according to Weinreich, Labov & Herzog 1968), a central tenet underlying the variationist framework, refers to the observation that recurrent linguistic choices made by a speaker in spontaneous discourse are not random. Rather, they are governed by various linguistic and social factors (Sankoff 1988). These constraints, which usually cannot be intuited and are not formally taught, constitute a portion of the grammar underlying the variability. In order to identify the underlying rules conditioning variant choice, *vernacular speech* is analyzed (i.e., unmonitored, informal, spontaneous speech). The vernacular is said to represent the "most systematic data for linguistic analysis" (Labov 1984:29) because it is one of the first varieties acquired and as such, it contains the least amount of style-shifting and hypercorrection (Tagliamonte 2006:8).

Once the variable context has been identified and the data have been extracted, each token is coded for a variety of linguistic and social factors that could potentially have an effect on variant choice. In order to identify such patterns, variationists often subject the data to both a distributional and multivariate analysis. A distributional analysis (i.e., a "factor by factor" analysis) provides key information about the correlation of individual factors with variant use (using percentages) (Tagliamonte 2006:193). A multivariate analysis, a key investigative tool in variationist sociolinguistics, allows the analyst to simultaneously test the effect of multiple factors on variant choice, using both percentages and probabilities (Tagliamonte 2006). In this study, *Goldvarb* is used to conduct the multivariate analyses, which allows for the identification of the rules underlying variant choice using three lines of evidence: the identity of the factors that contribute to variant choice in a statistically significant manner, the relative strength of these

factors (also referred to as the *magnitude of effect*), and the *direction of effect*, as instantiated in the constraint hierarchy (i.e., how the factors are ordered within the factor group) (Poplack & Tagliamonte 2001).

When conducting a multivariate analysis, the "orthogonality" (i.e., the independence) of the factor groups is paramount (Guy 1988:126–127 cited in Tagliamonte 2006:181). Goldvarb does not check for interactions between factor groups, therefore the onus falls on the analyst to control for the non-orthogonality of factor groups by manually checking for anomalies between the percentages and probabilities, and through cross-tabulation of the factor groups when necessary. If interactions are identified, they must be resolved before proceeding with the analysis in order to guarantee the validity and reliability of the results (Tagliamonte 2006).

Using these tools, a variationist analysis can, therefore, transcend the quantitative distribution of the variant (i.e., rates) by identifying the rules constraining the variability. Rates of occurrence do represent an important part of any quantitative analysis, but they can prove misleading in certain cases since they can be influenced by various extra-linguistic factors, including but not limited to style, prestige, who the speaker is addressing etc. (Poplack & Tagliamonte 2001:92). For this reason, sole reliance on rates of occurrence of a particular variant can lead to equivocal results. However, the analysis of rates *in conjunction* with the underlying linguistic and social constraints on variant choice can provide a fine-grained and detailed perspective on the variability at play.

2.2 The comparative method

Aspects of the comparative method are also adopted in this analysis. This method is of central importance in historical linguistics as it is commonly used for the purpose of linguistic reconstruction (Campbell 2013:108). This is achieved through the comparison of "forms from

related languages (*cognates*), which have descended from a common ancestral language (the *proto-language*) in order to postulate, that is to *reconstruct*, the form in the ancestral language" (Campbell 2013:110). This type of method is necessary for the reconstruction of the pre-settlement ancestor because traces of the language at this stage are usually rare, if not completely non-existent, all the more if the ancestor that is being reconstructed is a spoken vernacular – as is the case in this study. Spoken vernacular has been shown (e.g., Poplack & Dion 2009; Poplack, Lealess & Dion 2013) to differ dramatically from the written standard (which is what tends to be preserved in older texts).

This method operates under the principle that if a feature is found in two (or more languages) and this feature is not likely to have appeared accidentally or through other common paths (e.g., borrowing, etc.), then it is reasonable to assume that this feature was "transmitted from a common ancestral source" (Poplack & Tagliamonte 2001:96). Since it is unlikely that the presence of "systematically corresponding sounds" in different languages is nothing more than accidental (Campbell 1998:113), one can assume that the same would be true of underlying grammatical constraints (Poplack & Tagliamonte 2001:97). In other words, if varieties of a language share the internal conditioning of a particular variant, it is likely that these constraints were transmitted from a common source and did not arise separately in each variety (although this is also possible).

2.3 The comparative variationist method

By employing the comparative variationist method, it will be possible to establish, through a horizontal (synchronic) comparison of contemporary QF and contemporary HF, whether the variable use of specific morphosyntactic features was inherited from a common source or whether they are examples of linguistic innovation that occurred in one of the two

varieties. The results of both the distributional and multivariate analyses will reveal the degree to which the two varieties resemble one another in the areas of the grammar under analysis here. Both the rates of occurrence of specific linguistic features as well as the underlying constraints on the variants under examination are considered in the analysis. If the same factor groups are selected as significantly contributing to variant choice, the magnitude of effect of significant factor groups are ordered in the same way relative to the other factor groups, and the ordering of the factors within each factor group (i.e., the constraint hierarchy) are similar, this would suggest parallels in the rules underlying the variability across the two varieties and thus, would represent strong evidence to infer that this underlying grammar was inherited from a common source (i.e., HF). It would be highly unlikely that parallel underlying constraints (which are not formally taught and are almost always below the level of conscious awareness of speakers) and their direction of effect could have arisen separately in two different varieties that have little contact with one another. If the results reveal differences in the rules conditioning variant choice when comparing the two varieties, this would suggest that an innovation took place in one of them.

It should be noted that it may not be possible to ascertain which of the two varieties underwent innovation without comparison with a pre-colonization benchmark (i.e., HF from prior to the settlement of the French in Canada)[7]. A diachronic component (detailed below) is nonetheless included in this study with the hopes of identifying, if not the source of change, at least the existence of change in progress. However, because it spans only a few decades, it is limited in its explanatory power except in cases of rapid, recent change.

[7] The creation of a 17th century HF corpus of speech-like instantiations is a project that will be undertaken by Dr. Shana Poplack and colleagues at the Sociolinguistics Laboratory at the University of Ottawa, and this dissertation can be seen as a key component of that project.

2.3.1 Apparent vs. real time

Both apparent and real time analyses are methodological tools utilized by sociolinguists in order to identify potential linguistic change in progress. The first approach, apparent time, involves the analysis of specific linguistic variables according to age of the speakers. In other words, the analyst will observe the use of the variants across each age group and identify any potential patterns which point towards change in progress (Hockett 1950). One indication of a change in progress would be a monotonic pattern across age groups or simply a notable change from one group to another. However, such a finding does not always imply change. It could also suggest the presence of age-grading, which refers to change in linguistic behaviour affecting one specific age group which repeats itself across time (Hockett 1950).

The second approach, a real time analysis, involves observing the use of a linguistic variable at two specific points in time (Labov 1994:73). This method is generally preferred because a difference between two of the same age-groups at two different points in time cannot be considered the result of age-grading. It is usually indicative of change in progress. When two groups of different people are compared across time, this is referred to as a trend study (Labov 1994:76), which is the type of real time analysis employed here.

In the context of this study, both apparent and real time will be considered in order to determine, in the case of differences between HF and QF, which variety underwent the change. The real time span here is limited (25-40 years) and, therefore, will only allow for the identification of fast-moving or recent change in progress.

2.4 Selection of corpora

When studying the use of linguistic features in spoken language, it is important that naturalistic, unmonitored speech (i.e., the *vernacular*) form the basis of the analysis. One method of obtaining this type of speech is through a *sociolinguistic interview* (Labov 1984). Unlike typical interviews, one of the main goals of the sociolinguistic interview is to tap into the vernacular (Labov 1984). This is achieved through a variety of techniques designed to elicit unmonitored speech such as selecting appropriate questions as well as adapting questions to the age of the speaker (Labov 1984). Labov (1984:32–33) lays out several goals of the sociolinguistic interview, which include specific techniques that are meant to shift the style of the conversation toward the vernacular, for example, eliciting narratives of personal experience, stimulating group interaction (i.e., conversation that does not include the interviewer), as well as allowing the interviewee to take the lead in defining the topic of conversation (i.e., allowing the conversation to follow its natural course even if it is guided by specific themes).

In the case of the corpora that will form the basis of this study, it was important to determine how they were collected and to ensure that they consist of vernacular speech, especially given the fact that the QF corpora to which they will be compared contain highly informal, spontaneous speech (Poplack 1989; Poplack & Bourdages 2008). Since the vernacular is the main repository of these purportedly "QF" features, it is crucial to limit any comparison between the two varieties to their respective vernaculars. Below, the corpora used in the context of this study, as well as the steps that were taken in order to find HF corpora what were as equivalent as possible to the QF ones, are described in detail.

2.4.1 QF corpora

Two existing QF corpora are used in this study as a benchmark with which contemporary HF is compared: *Corpus du français parlé à Ottawa-Hull* (Poplack 1989) and *Le français en contexte: milieux scolaire et social* (Poplack & Bourdages 2008)[8]. Both of these corpora are housed at the Sociolinguistics Laboratory at the University of Ottawa and consist of informal, spontaneous speech.

The *Corpus du français parlé à Ottawa-Hull* (Poplack 1989) consists of 120 speakers who are stratified by age, sex, and neighbourhood. It includes speakers from both Ottawa (Ontario) and Hull (Quebec). It was collected in the early 1980s and constructed according to specific scientific sampling procedures. The only corpus of its kind, it was originally compiled to test for contact-induced change resulting from intense contact with English.

Le français en contexte: milieux scolaire et social (Poplack & Bourdages 2008) was compiled in 2005-2007 and is comprised of the spontaneous discourse of 190 speakers from a school in the neighbourhood of Montbleu (in Hull), which is one of the same neighbourhoods targeted in the *Corpus du français parlé à Ottawa-Hull* (Poplack 1989) approximately twenty-five years earlier. The speakers include students (all aged 16-18 years old) and teachers (ranging in age from 24 upwards). The motivation for the creation of this corpus was to compare the speech of students to that of their French teachers. Most analyses using these data usually compare and contrast these two groups.

[8] Poplack and associates have often included a third QF corpus in their diachronic analyses of several variables: *Les récits du français québécois d'autrefois*: *reflet du parler vernaculaire du XIX^e siècle* (Poplack & St-Amand 2009). Data from this corpus is not included in the initial comparisons with QF. Instead, the focus will rest mainly on the two corpora described above since they represent similar time frames to the HF corpora and thus are more suitable for comparison purposes. However, the 19th century QF corpus will be discussed briefly in Chapters 4 and 5.

Both of these corpora (Table 2.1.) have allowed for the diachronic analysis of many linguistic variables in QF by Poplack and her associates, and as such, will provide an ideal point of comparison for the analysis of similar variables in the HF data.

Corpus du français parlé à Ottawa-Hull			***Le français en contexte: milieux scolaire et social***			
Age	M	F		Age	M	F
15-24	10	10	Students	16-18	70	70
25-34	10	10	Teachers	24-34	1	7
35-44	10	10		35-44	0	5
45-54	10	10		45-54	2	0
55-64	9	10		55-64	0	5
65+	11	10		65+	-	-
TOTAL	**60**	**60**		**TOTAL**	**73**	**87**

Table 2.1: Distribution of speakers in two QF corpora[9],[10].

The QF corpora could be considered by some to not be representative of QF due to the fact that a) the 20th century corpus includes speakers on both sides of the Ontario/Quebec border (French is only the majority language on the Quebec side) and b) the level of contact with English is greater in this region than in any other part of Quebec. However, in addition to the fact that this area was settled by French colonists, it has been shown that the French spoken on either side of the border is virtually indistinguishable and that contact with English has not resulted in grammatical convergence for the variables that have been studied so far (e.g., Poplack 1997; Poplack & Levey 2010). Therefore, these corpora can be considered to be representative of mainstream Quebec French.

[9] These numbers represent the entire QF corpus. In their study of the subjunctive, Poplack, Lealess & Dion (2013) analyzed the speech of all of the 20th century speakers, but only that of the students in the 21st century. In her study of the expression of necessity, Lealess (2005) studied a subsample of twenty-two 20th century speakers, stratified by age and sex, while in this study, a subsample of twenty 21st century students, stratified by sex, is analyzed.

[10] Both corpora also account for various other external factors, including socioeconomic status and level of bilingualism in the 20th century one; social status, and stylistic register in the 21st century one. However, only age and sex are considered in this study.

2.4.2 HF corpora

As described above, large QF corpora, consisting of informal, spontaneous speech, have already been collected and analyzed by Poplack and her colleagues. In order to proceed with a comparative analysis between HF and QF, a first step involved finding comparable HF data, a task that turned out to be quite challenging for the reasons outlined below.

First, in order for the comparison to be meaningful, the nature of the data had to be comparable with the QF corpora (i.e., informal, spontaneous speech). However, many of the pre-existing HF corpora do not consist of vernacular speech. Instead, they contain written or scripted language or were collected in such a way that would inhibit unmonitored speech (e.g., formal interviews). The closer the data are to the (written) standard, the more the two varieties can be expected to resemble each other, albeit superficially. Moreover, if one were to compare formal HF to informal QF, this would amplify any potential differences between the two varieties by inflating the use of non-standard forms in QF and artificially lowering them in HF. In addition, some of the HF vernacular data previously reported on in the literature was not available for use by a third party.

A second issue concerned the size of the potential HF corpora. Morphosyntactic variation (the focus of this dissertation) tends to occur much less often in speech than other types of variation (e.g., phonological/phonetic). Therefore, a large quantity of data is often needed in order to undertake such an analysis. While some HF corpora (e.g., PFC: Durand, Laks & Lyche 2002) meet the first requirement (i.e., contain vernacular speech), the quantity of available data is simply not sufficient to complete a variationist analysis of morphosyntactic variation in most cases.

Additional challenges faced when attempting to find comparable HF corpora concerned the stratification of the speaker sample as well as the time frame in which the data were collected. Poplack's QF corpora were carefully assembled in the early 1980s and mid-2000s according to various levels of social stratification. It was important to find HF corpora that fell within this approximate time frame and that were stratified similarly in order to maximize comparability.

After careful consideration, two different HF corpora were chosen from which a large (> 1 million words) repertoire of pre-existing corpora of spoken HF discourse was assembled. The data originate from two different sources: *Corpus de Français Parlé Parisien des années 2000,* (Branca-Rosoff et al. 2012) and *Enquêtes Sociolinguistiques à Orléans* (Laboratoire ligérien de linguistique 1968), both of which are both publicly available. These two corpora contain informal spoken language in sufficient quantity to carry out this analysis. They are also socially stratified and were collected within similar timeframes as the QF corpora. The subsamples were carefully selected so as to parallel the socio-demographic characteristics (age and sex) of the QF corpora. A full description of each, as well as a justification of their informal nature, is presented below in Section 2.5

Some could argue that these two corpora are not sufficient as an HF baseline with which to compare QF speech since they are both from central France and the majority of settlers to *Nouvelle-France* originated from the western part of the country (Charbonneau 1990). While it would be ideal to analyze data from the exact regions from which the majority of New France settlers originated, these areas are now home to various regional dialects (Ager, 1990). In addition, as a result of globalization, Paris has virtually become "a melting-pot of French speakers from the whole country […]" (Blanchet & Armstrong 2006:257) and thus, due to the

increased the mobility of French speakers, dialect contact has undoubtedly played a role in the emergence of contemporary central HF vernacular speech. This is reminiscent of the situation in *Nouvelle-France* at the time of colonization. Crucially, it was most important to ensure that the data representing the HF benchmark be as comparable as possible to the QF corpora in terms of both quality and quantity.

2.5 Description of HF corpora

CFPP (*Corpus de Français Parlé Parisien des années 2000*) (Branca-Rosoff et al. 2012)

The first corpus used in the context of this study, *Corpus de Français Parlé Parisien des années 2000* (Branca-Rosoff et al. 2012), is a collection of informal Parisian French speech. With a population of approximately 2.2 million, Paris represents the largest urban centre of France (INSEE (Institut national de la statistique et des études économiques) 2010).

The CFPP consists of sociolinguistic interviews that were collected beginning in 2005-2006. By 2012, the corpus contained 535,000 words, 29 interviews, 44 speakers, and almost 38 hours of speech, all of which are fully transcribed. The transcription protocol for this dataset is very detailed as the researchers made a distinct effort to preserve the (mostly morphosyntactic and some phonological) variability found in natural speech (Branca-Rosoff et al. 2012). In addition, the transcribed interviews are aligned with their respective sound files, and are available online for consultation (Branca-Rosoff et al. 2012). One of the benefits of having the recordings aligned with the transcriptions is that it allows the researcher to return to the recording at any time in order to check any questionable occurrences or simply for verification purposes when required.

Nearly all of the speakers in the corpus were born and raised in Paris and are stratified according to various social factors, including age, level of education, neighbourhood, etc. Only

speakers born and raised in Paris for whom full demographic information was available were included in this study (Table 2.2).

Age	Male	Female	Total
15-24	5	3	8
25-34	3	3	6
35-44	2	1	3
45-54	1	3	4
55-64	5	5	10
65+	0	8	8
TOTAL	**16**	**23**	**39**

Table 2.2: Distribution of speakers by age and sex in the CFPP corpus used in this study.

ESLO (*Enquêtes Sociolinguistiques à Orléans*) (Laboratoire ligérien de linguistique 1968)

The second corpus, *Enquêtes Sociolinguistiques à Orléans*, contains two sub-corpora which span approximately forty years: ESLO1 (1969-70) and ESLO2 (2008-present). The speech contained in these two corpora comes from the city of Orléans, which is centrally located and about one hour's drive from Paris with a population of just over 100,000 (INSEE (Institut national de la statistique et des études économiques) 2009).

ESLO1 was originally collected in the late 1960s by British academics. It consists of various types of recordings, including public meetings, phone conversations, and face-to-face conversations/interviews. It is only the interviews that are of interest for this project since the focus is on variability in naturalistic speech. With more than 200 interviews and more than 300 hours of speech, it is considered to be the most extensive sample of spoken HF from the mid-20th century (Abouda & Baude 2006).

The creation of ESLO2, a continuation of ESLO1 approximately forty years later, was initiated by the CORAL (*Centre orléanais de recherche en anthropologie et linguistique*) with the purpose of compiling a corpus comparable to ESLO1 (Abouda & Baude 2006). As such, it

includes various speech types but like with ESLO1, only the interviews are of interest for this project.

ESLO1 (1968)				ESLO2 (2008-2011)			
Age	**Male**	**Female**	**Total**	**Age**	**Male**	**Female**	**Total**
19-24	5	7	12	20-24	0	1	1
25-34	7	7	14	25-34	3	3	6
35-44	5	3	8	35-44	2	3	5
45-54	6	3	9	45-54	2	3	5
55-64	6	8	14	55-64	1	1	2
65+	2	4	6	65+	3	5	8
TOTAL	**31**	**32**	**63**	**TOTAL**	**11**	**16**	**27**

Table 2.3: Distribution of speakers who were born and/or raised in Orléans by age and sex in ESLO1 and ESLO2 (all are included in this study).

The speaker sample in both ESLO corpora is quite varied (e.g., there is no requirement that the speaker be from Orléans to participate) and data collection is ongoing in the case of ESLO2. Speakers from both corpora are stratified by age, sex, and level of education. In this study, only speakers that were either born in Orléans or moved there before the age of 10 were included. Applying this criterion reduced the sample drastically. The sub-sample that will be used for this study is given in Table 2.3[11]. Speakers are fairly evenly stratified by age and sex in both corpora but in ESLO1, there are twice as many speakers that meet the criterion (born in Orléans or arrived before the age of 10) than in ESLO2.

Of particular interest in the ESLO corpora is the capacity for diachronic analysis across an approximate forty year period (1968-2008), which will also allow for as close a comparison as possible with two QF corpora. As noted in the previous section, if differences are revealed in the underlying grammar of a particular variant across the two varieties, this real time gap may provide clues as to which variety underwent the innovation, provided of course that it is either a

[11] There were two speakers that met the guidelines to be included in the sample whose interview transcripts were not available at the time of compilation of the subsample. Therefore, these speakers were excluded from the study. Also, the final versions of the transcripts were not available for ten of the speakers included in the sub-sample but I was granted permission to access the 'B' version (they are still waiting for their third and final correction).

change in progress or a recently completed change. If the change already reached completion prior to the 1960s, the real time span used in this study will not reveal which of the two is the innovative variety.

According to the gravity model of linguistic change (Trudgill 1974), the diffusion of linguistic change and innovation typically begins in large urban centres before moving to smaller cities and towns; therefore, it was important to ascertain at the outset whether there were differences between the two HF regions since it could be reasonably expected that Paris would be more advanced in the case of change in progress.

	Paris (CFPP)		**Orléans (ESLO2)**	
	N	**%**	**N**	**%**
Rate subjunctive	208/302	69	190/268	71

Table 2.4: Overall rate of subjunctive in HF (Paris and Orléans).

	Paris (CFPP)		**Orléans (ESLO2)**	
	%	**N**	**N**	**%**
falloir INF	180	47	209	50
falloir que	103	27	90	22
devoir	33	9	44	11
avoir besoin	8	2	13	3
être obligé	38	10	21	5
avoir à	24	6	39	9
TOTAL	**386**	**101**[12]	**416**	**100**

Table 2.5: Overall distribution of variants expressing necessity in HF (Paris and Orléans).

However, preliminary perusal of the data identified few linguistic differences between the two data sets for the variables studied here; therefore, they were combined to represent 21st

[12] Due to rounding.

century HF[13]. For example, as is detailed in Tables 2.4 and 2.5, the overall rate of subjunctive is nearly identical in the two regions as is the distribution of the variants used for the expression of necessity (both of these variables will be discussed in great detail in Chapters 3 and 4).

2.5.1 Access to the data

Both corpora (CFPP and ESLO) are (mostly) publicly available online. They were transcribed, catalogued, aligned, and annotated by their respective creators. Using the software *Concorder Pro*[14], all of the files in the subsamples of both corpora were concordanced in order to simplify the extraction processes that was used for the multiple variables analyzed in this dissertation. The concordancing process was fairly straightforward and simply involved changing the formatting in the original files such that they would be accepted and processed by the concorder. Concordancing the files facilitated the search for relevant *tokens* (i.e., occurrences of a particular variant) using a search function, which is the technique applied for all of the variables under investigation[15]. Word searches are carried out by locating a specific word (and all of its related forms) in the data. Then, the concordancing program lists each occurrence of the form in question along with its preceding and following context. The alternative to this facilitated extraction process is simply reading through the data in order to locate specific forms, which is not only much more time consuming, but often less accurate.

[13] As was pointed out by a member of the evaluation committee, some could argue that there are dialectal differences between Paris and Orléans and thus, they may not be comparable. However, when it comes to the morphosyntactic variables analyzed here, the two cohorts behave nearly identically. This does not exlude the possibility that differences could arise with other morphonsyntactic, phonetic, or phonological variables.

[14] *Concorder Pro* is a free online concordancing program that was modified by the Sociolinguistics Laboratory at the University of Ottawa.

[15] One of the disadvantages of this technique is that the concorder will only extract approximately 5-8 words on either side of the extracted word. It is, therefore, possible to miss some occurrences if they do not appear in this context. In order to address this issue, some contexts had to be checked manually.

2.5.2 Nature of the data

As mentioned earlier, the CFPP (Paris) corpus was collected with the objective of studying conversational Parisian French. Interviewers thus took various steps to encourage spontaneous speech within the sociolinguistic interviews (Branca-Rosoff et al. 2012). Each interviewer had a standardized list of topics to talk about (mainly about their experience in their neighbourhoods: the changes they have seen over time with new populations, etc.), but allowed the conversation to follow its natural course even if it meant deviating from the prescribed themes (Branca-Rosoff et al. 2012). A quick search of the CFPP data revealed the presence of swearwords, which are typical of informal speech (e.g., *putain, merde, foutre*, etc.), *tutoiement*, as well as many narratives of personal experiences. Also, many of the interviews were done in groups with at least two informants in addition to the interviewer, which is a context conducive to eliciting the vernacular. These facts suggest that these data contain informal speech and thus make them more comparable to the QF corpora considered in this study.

The ESLO (Orléans) corpora are not as easy to classify as informal speech. On one hand, ESLO1 was originally collected for didactic purposes – it was mainly meant to be used as a tool to help the British teach French as a second language in the British public education system (Abouda & Baude 2006). Within it, there is little swearing, a lot of *vouvoiement* (but this was also more common forty years ago), and fairly formal topics of conversation (these are listed in each interview file). For example, many speakers talk about the events of May 1968 (the massive revolts and strikes) and their opinions of the situation that led up to this uprising (unions, working conditions, etc.) since the interviews took place only a short time after these events (1969-1970). Some interviews focus mostly on informal topics (e.g., family, growing up, etc.) as well as many personal narratives. In addition, a lot of the interviews were done in groups, with

two or more individuals who were known to each other, thus facilitating the expression of vernacular speech.

ESLO2, collected since 2008, appears at first glance to be more informal than ESLO1. Unlike ESLO1, the list of topics discussed is not detailed in each file, but there are an abundance of personal narratives nonetheless. There are some swearwords in the data, but it is hard to tell just how informal it actually is or how closely it represents vernacular HF. This will be examined further using objective criteria to assess the level of formality in the following section. It is worth noting, however, that like with ESLO1 and CFPP, several of the interviews were done in groups of two or more and there is no shortage of personal narratives which suggests at least a certain degree of informality.

While this endeavour to match pre-existing HF corpora (collected by three different groups of researchers) with pre-existing QF corpora (collected by Poplack) was fraught with challenges, ultimately only data objectively deemed "comparable" was compared here. In other words, the nature of the HF data had to be assessed to confirm that it was, in fact, informal enough to be compared with the QF corpora. In order to achieve this, a stylistic analysis was undertaken, which is described below.

2.5.2.1 Stylistic analysis

Since the main purpose of choosing the HF corpora under investigation here was for comparability purposes, both with regards to the stratification and the nature of the data, additional steps were taken in order to ensure the informal nature of the HF speech. The informal nature of the QF data has already been established and discussed in detail in Poplack (1989) and Poplack & Bourdages (2008). In order to proceed, several measures of formality were applied to the HF data and then the results were compared with those found in another HF corpus,

identified as consisting of formal speech: ESTER (Galliano et al. 2005; cited in Torreira, Adda-Decker & Ernestus 2010), which contains journalistic recordings from several European and North African countries[16]. The data from ESTER are considered to be formal because they originate from planned and/or prepared speech intended for public radio broadcasting and thus represent a perfect benchmark against which to measure the level of formality of the two HF corpora under investigation here.

If there are differences between the two data sets, one would expect to find more non-standard linguistic behaviour in the HF corpora (comprised of sociolinguistic interviews) and more standard elements in ESTER (comprised of mostly scripted material). In Table 2.6, the overall frequency (per 1000 words) of hesitation markers and some common slang words in each of the four corpora is provided. Hesitation and slang words are prototypical markers of informal speech and tend to be avoided in formal contexts. The higher the rate of occurrence of such lexical items, the more informal the speech.

	ESLO1 (1968-1970) Freq. (raw N)	**ESLO2 (2008-2010) Freq. (raw N)**	**CFPP (2000) Freq. (raw N)**	**ESTER (1998-2004) Freq. (raw N)**
Hesitations				
Ben	2.48 (2045)	0.98 (395)	2.79 (1299)	0.26 (1766)
Euh	18.46 (15217)	28.79 (11639)	15.76 (7351)	8.14 (5452)
Hum	0.09 (77)	0 (2)	0.02 (9)	0.11 (73)
Slang words				
Merde	0.01 (6)	0.01 (5)	0.01 (5)	0 (1)
Putain	0 (0)	0.01 (5)	0 (2)	0 (0)
Cul	0 (1)	0.01 (5)	0 (0)	0 (0)
Con	0 (2)	0.02 (10)	0.01 (5)	0 (2)
Chier	0 (0)	0.03 (11)	0.01 (4)	0 (0)

Table 2.6: Frequency of certain hesitation and slang words in HF natural speech corpora that form the basis for this study compared with a journalistic HF corpus (ESTER results adapted from Torreira, Adda-Decker & Ernestus 2010).

[16] All of the results from ESTER originate from Torreira, Adda-Decker & Ernestus (2010) who used it as a benchmark against which to measure the informal nature of data contained in the Nijmegen Corpus.

The results clearly show that there is a higher use of both hesitation markers and slang words in the spoken HF corpora when compared with ESTER. The relative frequency of *ben* and *euh* is more at least double that of ESTER. While slang words are still quite infrequent in the HF spoken corpora, they are categorically absent from ESTER.

In addition to hesitation markers and slang words, the relative frequency of the negation marker *ne* was also compared across the four corpora. According to prescriptive norms, negative polarity is expressed in French by using *ne* along with a second negator (e.g., *pas, plus, rien,* etc.). An example is given in (1).

(1) « Moi, je ***ne*** crois ***pas*** que ce soit une différence de moyens. » (ESLO1/DM95/166)[17]

"Me, I don't think it's a difference of means."

However, *ne* omission has become quite common in informal spoken language, particularly in QF (e.g., Poplack 2015; Sankoff & Vincent 1980; Sankoff & Vincent 1977). Quantitative studies of *ne* omission in HF (e.g., Ashby 1981; 1976) suggest that its use is also quite rare, albeit not to the same degree as in QF. Its presence is also tightly correlated with level of formality: the more formal the speech, the more likely it is that *ne* will be retained (e.g., Poplack 2015).

Ashby (1976) looked the rate of *ne* amongst middle-class educated Parisians and found a much higher rate than with a socially diverse sample of speakers in 1981. Similarly, Moreau

[17] These and all subsequent examples are verbatim reproductions (including punctuation) from one of the three French corpora used in this study: *Corpus du Français Parlé Parisien des années 2000*-CFPP (Branca-Rosoff et al. 2012), and *Enquête Sociolinguistique à Orléans*, ESLO1 and ESLO 2 (LLL-Orléans 1968-2010). All examples from HF corpora are presented verbatim and thus reflect the transcription protocol adopted by the researchers. For the examples from the ESLO corpora, the speaker number reflects the one assigned by the corpus creators, but for the CFPP corpus, only speaker initials are provided since numbers were not assigned (see Appendix B for a list of the names of the speakers and their corresponding initials). All line numbers originate from the concordancing program that I used.

(1986) found a high rate of *ne* retention in television interviews, while Coveney (1996) found a much lower rate in a socially diverse corpus of spontaneous HF (Table 2.7).

	Ashby (1976)	**Ashby (1981)**	**Moreau (1986)**	**Coveney (1996)**
Type of data/speaker	Middle-class educated Parisians	Socially-diverse Parisians	Television interviews	Spontaneous speech
% *ne* retention	56%	37%	50%	19%

Table 2.7: Rate of *ne* retention in several HF studies**.**

In order further verify the level of (in)formality of the HF corpora that form the basis of this study, the rate of *ne* retention with *pas, rien, jamais*, and *aucun*[18] was examined. As is shown in Table 2.8, the rate of *ne* is much lower in the spoken corpora than in ESTER (the journalistic corpus) and the speakers from the 21st century spoken corpora exhibit an even lower rate of *ne* retention than those from the 20th century corpus.

	ESLO1 (1968-1970)	**ESLO2 (2008-2010)**	**CFPP (2000)**	**ESTER (1998-2004)**
% *ne* retention	32%	11%	13%	89%

Table 2.8: Rate of *ne* retention in CFPP, ESLO, and ESTER.

One might also suppose, based on the above results, that the 20th century HF corpus (ESLO1) is more formal than the 21st century ones (ESLO2, CFPP) due to the higher rate (nearly 3x) of *ne*. Ashby (2001) noted a clear decrease in the use of *ne* in all age groups in two studies with comparable subjects between 1976 and 1995 (Table 2.9), which suggests that this could be also be the case here.

[18] Only *rien, pas, aucun/e,* and *jamais* were considered here for comparability purposes. *Ne* can also occur with other negators like *personne, point,* and *plus.*

	Ashby (1976)	**Ashby (1995)**
% *ne* retention YOUNG	19%	14%
% *ne* retention OLD	52%	25%

Table 2.9: Rate of *ne* retention across time in HF in two samples of comparable speakers (adapted from Ashby 2001:6).

Although it is unknown if the 20th century HF corpus has a higher rate of *ne* due to stylistic differences or due to the decrease of *ne* across time, it is clear that the rate of *ne* in all three spoken HF corpora is much lower than what was found in ESTER. In addition, all three HF corpora show similar frequencies of hesitation markers and slang terms, which differ greatly from the rates found for ESTER, thus confirming their informal nature.

2.6 Comparison of QF and HF data samples

Table 2.10 shows the breakdown of the speaker sample (male and female speakers combined) in each corpus (both HF and QF) at two points in time: the 21st century (2005-2011) and the 20th century (1968-1981). Recall that both 21st century HF corpora (from Paris and Orléans, 2005-2011) were combined into one single subsample which will be compared with 21st century QF (2005-2006). The age-groups are largely based on those from the *Corpus du français parlé à Ottawa-Hull* (Poplack 1989). Despite differences in some of the cells, most of the age groups are well-represented in each corpus, which will enhance comparability.

20th C			21st C				
	QF Québec (1981)	**HF Orléans (1968)**		**QF Québec (2005-2006)**		**HF Orléans Paris (2005-2011)**	
Age	Total	Total		Age	Total	Age	Total
15-24	20	12	Students	16-18	140	15-24	9
25-34	20	14	Teachers	24-34	8	25-34	12
35-44	20	8		35-44	5	35-44	8
45-54	20	9		45-54	2	45-54	9
55-64	20	14		55-64	5	55-64	12
65+	20	6		65+	0	65+	16
TOTAL	**120**	**63**		**TOTAL**	**160**	**TOTAL**	**66**

Table 2.10: Comparison of HF corpora that will form basis of this study with QF corpora[19] that were previously analyzed (Period 1 and Period 2).

It is clear that the HF samples are much smaller than the QF ones, which means that this will significantly affect the amount of data available for analysis. However, the variationist method minimizes this discrepancy since rates of occurrence are proportional: not only are the contexts in which a particular variant occurs considered, but also those in which it could have but did not. Therefore, different corpus sizes should not affect the overall rate of a variant. In addition, this method focuses heavily on the underlying grammar, which should be readily discernible in most data sets, despite their size.

Summary

Now that the methodological and theoretical framework on which this study is based has been summarized, the ensuing chapters will focus on the analysis of three morphosyntactic variables (the subjunctive, the expression of necessity, and pronominal indefiniteness) in QF and HF in order to identify the (dis)similarities which may exist across the two contemporary varieties of French.

[19] Recall that only a subsample of the QF corpora was considered in most of the QF studies with which the findings of this study will be compared.

Chapter 3 THE SUBJUNCTIVE

3.1 Introduction

Tense, aspect, and modality all form part of the grammatical system of any language. Tense and aspect refer to the time and nature of the event respectively, while modality is "concerned with the status of the proposition that describes the event" (Palmer 2001:1). One of the ways that modality can be expressed is through the use of verbal mood, which can be defined "the difference between clauses which is marked by indicative or subjunctive verb forms in languages which are traditionally described as having an opposition between such forms" (Portner 2011:1262). In French, grammarians report there to be four distinct verbal moods: the indicative, the conditional (although this is somewhat controversial), the imperative, and finally, the subjunctive, which is the focus of the current chapter.

The subjunctive is a hot topic for grammarians and theoreticians alike, who exhibit little agreement on the contexts in which it is, or *should* be used. There is also widespread disagreement on its status in the language: on one end are those who insist that its use is dwindling (e.g., Dreer 2007; Loengarov 2006; van der Molen 1923; Bauche 1929), while on the other end are those who assert that it is alive and well (e.g., Barral 1980; Glatigny 1976; Cohen 1965). In addition, much of the literature that discusses the subjunctive revolves around the question of the degree to which its use is semantically motivated.

However, despite such disagreement and inconsistencies across sources (which will be discussed in more detail below), there are certain facts upon which most linguists and grammarians agree. First, it is generally agreed that the subjunctive is "*un mode de la*

dépendance" (‘a dependent mood’) (Glatigny 1976:19). The subjunctive has become exceedingly rare in main clauses (Harris 1978:171), where it is usually limited to fixed expressions (e.g., *vive l’Europe* ‘long live Europe’) and in many cases, main clause subjunctives are preceded by the complementizer *que*, a marker of subordination (e.g., *qu’il vienne…* ‘let him come…’). Harris (1978:172) suggests that main clause subjunctives are “residual survivors of an earlier stage in the language’s development”, referring to their transfer from Latin, where they were one of the few contexts in which the subjunctive was said to be semantically motivated.

In contemporary French, the subjunctive is largely limited to dependent or subordinate clauses, a context which became associated with the subjunctive in Latin at a later stage and was subsequently transferred to French (Harris 1978). Subjunctive-selecting subordinate clauses are headed by a) a verbal governor (e.g., *falloir* ‘to be necessary’, *vouloir* ‘to want’, etc.) or b) a non-verbal governor, which includes various adverbial conjunctions (e.g., *pour que* ‘so that’, *bien que* ‘although’, etc.). An example of a verb carrying subjunctive morphology embedded under a verbal governor (*falloir*) is given in (1), while an example under a non-verbal governor (*pour que*) is given in (2).

(1) « Pour un métier comme ça, faut que ça se ***fasse*** (S) naturellement. » (ESL02/9/114)
‘For a job like that, it should come naturally.’

(2) « Je bats, je mets du lait pour que ça ***puisse*** (S) gonfler. » (ESLO2/MF363/1566)
‘I beat it, I add some milk so that it can become fluffy.’

While most agree that the subjunctive alternates with the indicative and the conditional in many contexts that are purported to be subjunctive-selecting, they do not agree on the reasons as to why this occurs. Some consider such examples to be reflective of inherent variability (e.g., Poplack, Lealess & Dion 2013; Poplack 1992; 1990), while others suggest that there are

differences in the level of assertion or the attitude of the speaker in each of them, which has led to differences in the inflectional morphology assigned to the embedded verb (e.g., Dreer 2007; Imbs 1953; van der Molen 1923). In other words, many argue that the use of the subjunctive is semantically motivated.

The variable nature of the subjunctive is illustrated below with an example of an indicative (3), a subjunctive (4), and a conditional (5) under the same verbal governor (*croire* (neg.) 'to not believe'). In addition, both a subjunctive and indicative can be found in the same utterance, with the same verb, embedded under the same verbal governor, as in (6), which casts serious doubts on the idea that the use of the subjunctive is semantically motivated.

(3) « Si l'enfant réussit ou réussit pas je crois pas que c'***est*** de la faute du professeur.» (ESLO/SD272/145)

'Whether the student is successful or not, I don't think that it's the teacher's fault'

(4) « Je ne crois pas que ce ***soit*** d'une grande importance. » (ESLO1/TS165/978)

'I do not think that it is of great importance.'

(5) « Je crois pas que je ***prendrais*** (C) des dispositions spéciales (ESLO1/MD461/760)

'I do not think I would make special arrangements.'

(6) «…Il faut qu' tu ***fais*** (I) ci faut qu' tu ***fasses*** (S) ça» (CFPP/AH/4632)

'You have to do this, you have to do that.'

Note that there are several contexts in which it is impossible to identify if the embedded form is a subjunctive or an indicative, including with all first group verbs (i.e., verbs with an *-er* ending). Table 3.1 highlights ambiguous first group forms.

***aimer* ('to like')**			
	Present Subjunctive	**Present Indicative**	**Imperfect Indicative**[20]
1st sg. *je*	*aime*	*aime*	*aimais*
2nd sg. *tu*	*aimes*	*aimes*	*aimais*
3rd sg. *il/elle*	*aime*	*aime*	*aimait*
1st pl. *nous*	*aimions*	*aimons*	*aimions*
2nd pl. *vous*	*aimiez*	*aimez*	*aimiez*
3rd pl. *ils/elles*	*aiment*	*aiment*	*aimaient*

Table 3.1: Conjugation of a typical first group (*-er*) verb in the imperfect, present indicative, and present subjunctive (shading indicates ambiguous forms).

The same issue arises with some non-first-group verbs as well, as is illustrated in Table 3.2 for *finir* 'to finish', where ambiguous forms are once again highlighted. All ambiguous forms must necessarily be excluded from any study of the subjunctive.

***finir* (to 'finish')**			
	Present Subjunctive	**Present Indicative**	**Imperfect Indicative**
1st sg. *je*	*finisse*	*finis*	*finissais*
2nd sg. *tu*	*finisse*	*finis*	*finissais*
3rd sg. *il/elle*	*finisse*	*finit*	*finissait*
1st pl. *nous*	*finissions*	*finissons*	*finissions*
2nd pl. *vous*	*finissiez*	*finissez*	*finissiez*
3rd pl. *ils/elles*	*finissent*	*finissent*	*finissaient*

Table 3.2: Conjugation of *finir* verb in the imperfect, present indicative, and present subjunctive (shading indicates ambiguous forms).

The second point upon which most researchers agree is that the only remaining forms of the subjunctive used in present-day French speech (and for the most part in writing) are the present and present perfect forms (e.g., for *pouvoir* 'to be able': *il puisse, il ait pu*). The imperfect and pluperfect subjunctive (e.g., for *pouvoir* 'to be able': *il pût, il eût pu*) have basically disappeared from the language and are only found sparingly in literary texts (Grevisse & Goosse 2014; Glatigny 1976).

[20] In some studies, 1st and 2nd person plural forms are included because it can be argued that, in most contexts, it is possible to discern whether it is a subjunctive or an imperfect. However, in order to maintain complete objectivity, these (rare) forms were excluded from the present analysis.

This chapter consists of a variationist analysis of subjunctive use in 20th and 21st century HF under both verbal and non-verbal governors, complemented by a thorough comparison with its use in QF (based on the results of Poplack, Lealess & Dion 2013). Since it is an example of a linguistic feature whose use (or lack thereof) is generally *below* the level of conscious awareness of speakers, its absence is not overtly stigmatized, at least not in QF (Poplack 2015:300)[21]. Before proceeding to the results of the analysis, a review some of the relevant literature, along with a description of the methodological approach employed in this study are presented below.

3.2 Background literature

3.2.1 Semantic considerations

The abundance of literature focusing on the subjunctive in French includes prescriptive, descriptive, theoretical, and empirical accounts of how, where, and why it is used. Many researchers list all of the subjunctive occurrences they can find in various sources and attempt to classify them into semantically motivated categories where all of its members behave near-identically (e.g., Nordahl 1969), while others make use of such categories established elsewhere in the (usually prescriptive; see Grevisse & Goosse 2008; 1989) literature in order to check for subjunctive use in a particular data set (e.g., Sand 2003; 1983).

Limiting oneself to prescriptively defined contexts in any study of the subjunctive or any variable for which such injunctions have been assigned can be problematic when one considers the variety and volatility of grammarians' prescriptions. With regards to the subjunctive in particular, Poplack, Lealess & Dion's (2013) review of more than 160 grammars across

[21] Poplack (2015) found that French teachers did not use more subjunctive than their students, and that students actually used fewer subjunctives in a formal context than in an informal one. Since social conditioning represents an important stigmatization cue, these findings provide support for the notion that subjunctive use generally falls below the level of conscious awareness of the speakers.

approximately five centuries revealed that more than 600 subjunctive governors had been prescribed over that period of time and that the number continues to rise. Moreover, there was great inconsistency and contradiction in the governing elements that were prescribed to co-occur with the subjunctive. The same type of disjointedness can be found in linguistic research: semantic categories identified as requiring the subjunctive tend to vary from one study to the next as one researcher after another attempts to account for all of the contexts in which the subjunctive is (or *should* be) found.

As mentioned above, these persistent and extended disagreements amongst linguists and grammarians mostly stem from whether or not subjunctive use is semantically motivated (Ayres-Bennett, Carruthers & Temple 2001). Many claim that the choice of using the indicative or the subjunctive is meaningful and that the meaning is largely "dependent on the degree to which the speaker wishes to commit herself to the reality or truth value of the complement proposition" (Poplack, Lealess & Dion 2013:140).

The realis/irrealis distinction is one that has often been used to distinguish contexts that require the subjunctive from those that require the indicative. Realis refers to a situation which has been 'actualized' (i.e., has occurred or is in progress), while irrealis refers to those which have not been and remain "in the realm of thought" (Palmer 2001:1). However, this distinction is weakened by the fact that the subjunctive can appear in typical realis contexts, i.e., ones in which the event has already taken place (e.g., *bien qu'il* ***soit*** (S) *venu* 'although he came') and the indicative can be found in irrealis contexts, i.e., those which have not yet occurred (e.g., *je pense qu'il* ***vient*** (I) 'I think that he is coming') (Ayres-Bennett, Carruthers & Temple 2001:195)

Some argue that distinction between the indicative and the subjunctive can be better represented using the notions of assertion and non-assertion (e.g., Hooper 1975 cited in Palmer

2001:3). Rather than depending on whether or not the utterance is factual, this approach depends on the degree to which it is asserted (Palmer 2001:4). However, after testing the meaningfulness of this distinction in QF, Poplack (1992:238) found that in each of categories of degrees of assertion proposed by Hooper, there was clear variability in subjunctive use with no preference for assertive or non-assertive contexts[22].

There are a host of other semantic meanings that have frequently been linked to subjunctive use. One is that it expresses "possibility as opposed to probability" (Ayres-Bennett, Carruthers & Temple 2001:195). Guillaume (1929:32–36 cited in Ayres-Bennett, Carruthers & Temple 2001:195) states that this distinction can be accounted for by the fact that when something is possible, it is not necessarily probable, but when something is probable, it is most definitely possible. However, just like with the realis/irrealis and assertion/non-assertion distinction, counter-examples cast doubt on this classification. For example, Martin (1983:108–110) and Dreer (2007:xx) both give examples of subjunctive use with *le fait que* 'the fact that', despite it clearly expressing a fact, as opposed to a possibility.

Another commonly-cited association is that of the subjunctive with contexts that mark doubt or subjectivity. But it is not uncommon to find examples of the subjunctive in cases which clearly do not express doubt (e.g., *Je ne doute pas que X* "I don't doubt that X') (Ayres-Bennett, Carruthers & Temple 2001:196). The subjunctive has also been referred to as "*le mode de l'énergie psychique*" 'psychic energy'" (Le Bidois & Le Bidois 1935:501), which suggests that some type of psychological motivation is behind each use of this form.

The semantic nuances associated with the subjunctive outlined above represent only a small portion of those that have been invoked in the literature. Others include the attitude of the

[22] A comprehensive account of the realis/irrealis and assertion/non-assertion distinctions can be found in Poplack, Lealess & Dion (2013), Palmer (2001), Poplack (1992), and Hooper (1975).

speaker (van der Molen 1923), the degree to which a speaker commits to what is being expressed by the utterance (Dreer 2007), or the notion of futurity (Lagerqvist 2003), amongst others.

One of the problems with the attempts to semantically differentiate the subjunctive from the indicative is that most of the meanings are nearly impossible to identify in running discourse. The feelings and semantic nuances which are said to be required for subjunctive use are often not discernable in speech by the speaker or the listener (Poplack, Lealess & Dion 2013; Poplack 1992; Glatigny 1976) – much less the analyst. In other words, one cannot possibly know with any degree of certainty the sentiment(s) and intention(s) of a speaker during a speech act without a direct window into the mind of that individual unless there are other contextual indicators.

Some argue that subjunctive use in contemporary French is "an empty, formal marker devoid of semantic content" required by the main verb (Ayres-Bennett, Carruthers & Temple 2001:197). One researcher who adopts a similar approach is Gross (1975; 1978), who rejects the semantic function of the subjunctive and instead suggests that the only difference between the subjunctive and the indicative is morphological. Foulet (1919) proposes that subjunctive use is mostly governed by style and the use of the subjunctive over the indicative is not due to a change in meaning, while Brunot (1922:522) suggests that its use is the result of "*la servitude grammaticale*" 'grammatical servitude'[23],[24].

Harris (1978) not only suggests that the use of the subjunctive contains little, if any, semantic content in contemporary French, but that this pattern reflects a much earlier stage in the language. He refers back to its use in Latin, the source variety, where it was extended from

[23] See Poplack, Lealess & Dion (2013) and Ayres-Bennett, Caruthers & Temple (2001) for an overview of the semantic classifications commonly assigned to the subjunctive mood as well as a detailed summary of how its use has been approached theoretically and empirically in the literature.

[24] In some recent semantic literature, it is acknowledged that the subjunctive can alternate with other moods in certain contexts in French without resulting in a change in meaning, and that even when semantic movivations are at play, they "may be blurred by other factors" Godard (2012:147). For an overview of formal semantic approaches to verbal mood, see Portner (2011) and Godard (2012).

semantically motivated main clause contexts to subordinate clauses, primarily as an additional marker of subordination, which resulted in a "purely formal change required by the grammatical rule of the language" (Harris 1978:169, 173), and subsequently spread to contexts which would have never elicited a subjunctive. According to Harris (1978:174), its use in contemporary French can be characterized as a "meaningless marker of subordination in certain contexts", which are often structurally or lexically defined.

Variationist studies of the subjunctive in QF (e.g., Poplack, Lealess & Dion 2013; Poplack 1992) have operationalized some of the main semantic claims in order to quantitatively test if they are, in fact, conditioning subjunctive use. The results have consistently shown that in any case where there appears to be a semantic effect, an underlying lexical effect is responsible, and that the choice between the subjunctive and the indicative is actually due to lexical or structural factors as opposed to semantic ones. These studies are discussed in more detail below.

In the present study, both semantic (as operationalized by Poplack, Lealess & Dion 2013) and morphosyntactic conditioning factors are tested in order to assess their relative effects on subjunctive use in HF and the results are then compared to what was found for QF. These will be explained in more detail in the ensuing sections. Even though Poplack found that the subjunctive had no meaning in QF, the inference that could be drawn is that it lost its meaning at some point as the result of change. In this connection, it will be particularly important to ascertain whether the subjunctive has retained its meaning (if it ever had one to begin with) in the source variety (i.e., HF), which is, in part, one of the goals of the present study.

3.2.2 Status of the subjunctive in present-day HF

Even though it is discussed at length in both descriptive and empirical studies, the status of the subjunctive in 20th and 21st century HF remains a point of contention amongst researchers. Some argue that it is losing ground to the indicative (e.g., Dreer 2007; Laurier 1989; Bauche 1929; van der Molen 1923), while others assert that it is used quite regularly (e.g., Blanche-Benveniste et al. 1990; Nerio 1979; 1978; Cohen 1965). Some of those who assert that the subjunctive is declining also suggest that the contexts in which it can be used are also decreasing and as a result, it is being replaced by infinitival structures (e.g., Bauche 1929; Dreer 2007:214), the future indicative, or the conditional (e.g., Lagerqvist 2003; Togeby 1974; 1966). Togeby also suggests (citing Boysen 1971 and Börjeson 1966) that the *passé simple* and the imperfect are used in place of the subjunctive to express temporal references which can no longer be expressed by the subjunctive uses that remain in contemporary French[25].

Interestingly, of all of the studies cited here that take a stance on the status of the subjunctive in contemporary HF, none (with the exception of Dreer 2007) are diachronic. In fact, some are not even empirical. In other words, most of the researchers have interpreted variability between the subjunctive and the indicative as evidence of the decline of the subjunctive[26]. For example, van der Molen (1923) makes strong claims about the disappearance of the subjunctive and, while he admits that it is still much more frequent in French than in English, he suggests that its disappearance is part of a more general simplification trend in Romance languages. His assertions are based on his descriptive overview of subjunctive use in letters from French

[25] It should be noted that the *passé simple* is rarely used in contemporary French speech.
[26] See Auger (1990) for a summary of the findings of several descriptive studies of the subjunctive in HF. Only discuss select empirical ones are discussed in detail here.

soldiers, but he does not quantify his findings nor does he attempt to explain the variability other than to invoke the idea that there is a psychological reason behind its use.

3.2.2.1 Quantitative studies of the subjunctive in HF

There is no shortage of studies examining the subjunctive in HF descriptively, theoretically (e.g., Abouda 2002; Lalaire 1998; Huot 1986; Togeby 1974; Togeby 1966; Imbs 1953), and quantitatively. Below, several quantitative studies of subjunctive use in HF are summarized in order to situate their findings with regards to ours. They are presented chronologically with the intention of offering a glimpse into the evolution of the study of the subjunctive in HF across time.

Gougenheim et al. (1964)

This study represents one of the first quantitative studies of subjunctive use in HF. Gougenheim et al. analyzed the use of the subjunctive in the spoken language of 275 speakers who originated from all over France and from various other French-speaking countries. The authors extracted all unambiguous (i.e., contexts in which the subjunctive form differs from the indicative) and ambiguous (i.e., contexts in which the surface form of the subjunctive and the indicative are identical) subjunctive forms and noted the verbal or non-verbal governor of the main clause. However, they only provide raw numbers of subjunctive and ambiguous occurrences without mention of the indicative, which makes it impossible to determine at what rate the subjunctive is employed in most of the contexts. Recall that, as was illustrated in Section 3.1, ambiguous forms refer to cases where the morphological form of the embedded verb is such that it is impossible to tell if it is a subjunctive or an indicative since the surface forms are identical.

While the description of the findings is quite limited, the authors do note that *falloir* and *pour que* co-occur with the subjunctive more often than the other verbal governors (Gougenheim et al. 1964:213). The other non-frequent governors include *vouloir* 'to want', *dire* 'to say', *attendre* 'to wait', *être +attribute* 'to be +attribute', *avoir peur* 'to be scared', *avant que* 'before', *à moins que* 'at least', and *bien que* 'even though' (Gougenheim et al. 1964:213). No general conclusions with regards to the status of the subjunctive or its semantic standing can be drawn from the findings.

Nordahl (1969)

Also among the first quantitative studies of subjunctive use in HF, Nordahl analyzed more than 24,000 contexts from written sources, including plays, newspapers, novels, and various other literary works (ibid., 10). He extracted all nominal subordinate clauses that "were of interest"; however, he does not expand on exactly how such cases were defined except to say he focused on (mostly) unambiguous examples of mood choice (ibid., 10). He classified subjunctive governors into three distinct classes: volitive, subjective, and dubitative. The dubative class consists of ten subgroups divided into two classes that lack a cohesive definition (according to him). Within the above-mentioned classes, he also distinguished between *contactuel* (i.e., no distance between the governor and the embedded verb), locutional (i.e., *ce que* 'that', *de ce que* 'of that'), attributative (i.e., *le/la* N *est que*… 'the N is that'), emphatic (i.e., *ce qui est* x *c'est que*), and *emphatisante* (i.e., *le/la* N *c'est que* 'the N it's that') contexts. In addition, he considered the *valeur énonciative* 'enunciative value' of the sentence (interrogative, affirmative, negative).

Nordahl found variable subjunctive use in most of the semantically-defined contexts he proposed. According to him, the subjunctive is obligatory with volitive verbs and he attributes

counter-examples to "*un glissement sémantique"* 'a semantic slip' (ibid., 44, 248), which refers to the fact that in certain contexts, the semantic nuance expressed by the governor changes from volition to declaration when an indicative is used (ibid., 44). In subjective contexts, he found the subjunctive to only be obligatory in contexts devoid of intervening material, and found no effect based on the enunciative value of the sentence. His results for the dubitative class were varied as not all of the subgroups behaved in the same way.

While Nordahl does provide subjunctive rates for many individual governors, he does not group them together in order to determine the rate of the subjunctive overall or even across all of his established categories. Additionally, Auger (1990:11) notes that his placement of verbs into the different categories is sometimes questionable (e.g., *s'habituer* 'to get used to' as a volition verb) as well as the fact that, at times, the same verb belongs to more than one category (e.g., *expliquer* 'to explain' is included under volition and dubitative verbs). Overall, this study adds little to what was known at the time about the use of the subjunctive.

Bieler (1972)

Bieler analyzed subjunctive use in both written and spoken language. Written language was represented by a collection of French newspapers, while spoken language was represented by both casual (i.e., naturalistic speech recorded using methods similar to those used in a sociolinguistic interview) and directed interviews (i.e., where the speaker was asked to complete sentences which were found to be strongly associated with the subjunctive in the other two data sets). He found little difference between the casual interviews and the newspaper data in terms of subjunctive use: 5.8% and 5.3% respectively. However, it should be noted that these figures are based on the total number of subjunctives used out of the total number of embedded clauses. In other words, this includes embedded clauses that have never been prescriptively or descriptively

associated with subjunctive use since the subjunctive is not prescribed in all embedded clauses. With regards to the interviews in which speakers were asked to complete various sentences with an embedded verb (all of which were contexts strongly associated with the subjunctive (n= >5) in the unguided interviews and newspaper analysis), a subjunctive was picked 2676 times out of a possible 7080 choices (38%).

He concluded that subjunctive use is more frequent and stable in specific (semantically-defined) contexts, such as after verbs of desire, emotion, possibility, necessity, and doubt, but also suggests that the overwhelming stability of the subjunctive in these contexts likely signals "*la servitude grammaticale*" 'grammatical servitude' (ibid., 60). He found the subjunctive to be less frequent and less stable in contexts in which the speaker had the choice to differentiate between two different meanings (e.g., realis vs. irrealis) but conceded that it was not clear if the speaker was actively making the choice of one variant over another to convey a given meaning or if they were simply unaware that the choice existed. Finally, despite asserting that the subjunctive is abundant in contemporary French, he hinted at the possibility of a decline in its use, based on the fact that in apparent time, young speakers used less of it than older speakers, a result which could also be attributable to age-grading.

Nerio (1979; 1978)

Nerio studied the use of the subjunctive in spoken HF in his doctoral thesis (1978) and later in an article summarizing the main findings (1979). His corpus of spoken language consisted of radio and television interviews as well as recorded conversations but by his own admission, the corpus did not contain enough vernacular speech (1979:36). His aim was to find a correlation between the different tenses used in prescriptively motivated subjunctive-selecting contexts and various social and linguistic factors. He concluded that the subjunctive is far from

disappearing (with an overall rate of occurrence across all contexts of approximately 85%). His analysis of linguistic factors included the semantic contexts proposed by Nordahl (1969) as well as the tense of the embedded verb as discussed by Togeby (detailed below). He also confirmed Nordahl's claim that the indicative is quite common in *emphatisante* and emphatic contexts, but did not have enough data to verify its use in Nordahl's other two '*indicativogènes*' contexts (i.e., contexts said to favour the indicative: attributive and locutional).

According to Nerio, approximately half of the indicative occurrences could be accounted for by the fact that other tenses were used in place of the subjunctive in order to communicate temporal nuances that the subjunctive is not capable of expressing. According to Togeby (1966; 1974), the present and past subjunctive are incapable of expressing future temporal reference; therefore, speakers resort to using the future or conditional in such instances. Boysen (1971) and Börjeson (1966) suggest that speakers often resort to using the imperfect and the *passé simple* to express perfect and imperfect temporal reference.

In short, despite suggesting that subjunctive use is indeed abundant, it is clear that there is variability in certain contexts. However, his analysis does not provide any clues as to why one tense is chosen over another in embedded clauses capable of hosting the subjunctive, nor does he shed any new light on why the subjunctive is used categorically in some contexts and variably in others. In fact, he doesn't discuss contexts in which the subjunctive is categorical at all, as is explained by Auger (1990:14).

Vandergheynst (1982)

Vandergheynst analyzed mood and tense concordance between the main clause and subordinate clauses introduced by *que* 'that' in spoken HF (including non-subjunctive-selecting ones). He used a subsample of the ESLO1 corpus (also used in this study) and extracted all

subordinate clauses introduced by *que*. Similar to many other studies of the subjunctive, he categorized the contexts under analysis into various semantic classes. While he did distinguish between the subjunctive, the indicative, and ambiguous embedded verbs, he concluded in most cases that the ambiguous forms were actually subjunctives. This assumption was based on the lack of unambiguous indicative embedded verbs found in those contexts. In cases where a subjunctive was expected but not used, he argued that the use of the indicative was semantically motivated or the result of the need to express a temporal reference incapable of being expressed by the subjunctive. He also noted, in at least one case, the role of the distance between the governor and the indicative embedded verb, stating that the intervening material caused the speaker to forget that they 'should' have used the subjunctive.

He found the subjunctive to be nearly categorical with verbal governors expressing volition, but practically absent from those expressing an opinion or perception. In his analysis, *falloir que* and *pour que* emerged not only as two of the most frequent governors, but also as having very strong associations with the subjunctive, although this is never explicitly stated.

His general conclusions with regards to the use of the subjunctive in spoken HF were twofold: not only do speakers use indicatives where a subjunctive should be expected (from a prescriptive point-of-view), they also avoid the use of contexts thought to be associated with the subjunctive more generally. It should be noted, however, that the number of tokens upon which his conclusions are based is quite limited in many contexts. For example, the entire class of volitive verbal governors consists of twenty-one unambiguous embedded verbs, while the emotive verbal governor category consists of only six.

Blanche-Benveniste et al. (1990)

Blanche-Benveniste et al. (1990) present the results obtained by Lim (1989) in her study of subjunctive use in mainstream HF using the GARS (*Groupe aixois de recherche en syntaxe*) corpus of spoken HF (Blanche-Benveniste & Jeanjean 1987). Only the unambiguous subjunctive occurrences were extracted (n= 452) and their structural context was noted. In other words, non-applications were not extracted so no overall rates of use were provided. The distribution of subjunctive morphology in the data is provided in Table 3.3.

%	Type of construction
35	Impersonal constructions (e.g., *il faut que* 'it is necessary that', *c'est important que* 'it is important that', etc.)
28	Personal verbal constructions (e.g., *je veux que* 'I want that', *j'aimerais que* 'I would like that', *j'ai peur que* 'I'm scared that', etc.)
21	*Pour que* 'so that', *pour pas que* 'so not that', *avant que*, 'before' *à moins que* 'unless', *quoi que* 'whatever'.
6	Alternatives (e.g., *que ce soit l'un ou l'autre* 'that it be either one or the other', *que je réussisse ou rate mes examens* 'that I pass or fail my exams')
4	Relatives, usually with *pouvoir* 'to be able' and an indefinite subject (e.g., *je n'ai pas de personnel qui puisse faire ça* 'I don't have any staff that are able to do that')
6	Other (e.g., optatives, *le fait que*… 'the fact that… ')

Table 3.3: Subjunctive morphology according to structural context in Lim (1989) (adapted from Blanche-Benveniste et al. 1990).

Lim's results showed that, while certain constructions (e.g., *le fait que* 'the fact that') and fixed expressions (e.g., *qu'il vienne ou qu'il vienne pas* 'whether he comes or not', *quoi que ce soit* 'whatever it may be') were used more by those with a higher social status, subjunctive use was present among all of the speakers in the sample, but they did not specify at what rate. Since

they did not account for any non-applications (i.e., places where the subjunctive could have occurred but did not), this study gives no information as to how productive subjunctive use actually is in this data set. Blanche-Benveniste et al. do not make any direct claims as to the productivity or semantic contribution of the subjunctive, but rather provide a descriptive account of its use according to Lim's findings.

Chauveau (1998)

Chauveau (1998) compared mood choice in two different North-American communities in Canada to its use in Haute-Bretagne, France. For the three communities, the data consisted of examples appearing in lexical and regional dictionaries. He found a much higher rate of subjunctive in the Canadian communities (51% in Port-au-Port in Western Newfoundland; 66% in Saint-Pierre and Miquelon, which is located just off the coast of Newfoundland) than in Haute-Bretagne, where the rate of subjunctive was only 5%. These rates are based on a limited number of tokens (n= 200 total for both Canadian communities and n= 43 in Haute-Bretagne) that were extracted from the dictionaries.

He concluded that subjunctive use in many contexts is no longer a choice but rather that it is strictly constrained by the structural context (i.e., the governor), and is, therefore, semantically redundant. Perhaps most surprising is his finding of a rate of 5% subjunctive in the HF sample as it is *much* lower than expected, and than what has been reported elsewhere. This can likely be explained by the methodology he employed: extracting sample sentences in subjunctive-selecting contexts from regional dictionaries can hardly be expected to produce results that are representative of the community norm.

Lagerqvist (2003)

Lagerqvist (2003) analyzed subjunctive use in a newspaper corpus of HF. Each potential subjunctive-selecting context examined was semantically categorized by the author, but the categorization was never justified (i.e., he doesn't explain why he only looked at specific governors that fall under specific pre-determined semantic categories, but it is safe to assume that it was prescriptively motivated). In Table 3.4, the distribution of subjunctive occurrences in his study according to syntactic classification is shown.

%	Type of construction
11	Main clause (e.g., *que Dieu* ***soit*** *bien* 'that God is good', *que tu te* ***taises*** *(S)* 'that you be quiet', etc.
50	Complement clauses – subject, object (e.g., *je crains que* 'I fear that', *la crainte que* 'the fear that')
14	Relatives – restrictive and concessive (e.g., *je cherche une femme qui* ***ait*** *(S) des cheveux blondes* 'I am looking for a woman who has blonde hair'*; quoi qu'il* ***fasse*** **'**whatever he does')
25	Subordinate clause introduced by an adverbial governor

Table 3.4: Subjunctive morphology according to structural context in Lagerqvist (2003) (adapted from Lagerqvist 2003:154).

Lagerqvist classified the subordinate clauses into four different semantic categories: those governed by a volitional verb (e.g., *souhaiter* 'to wish', *demander* 'to ask', *vouloir* 'to want', *attendre* 'to wait'), by a verb expressing sentiment, and impersonal expressions (*il faut* 'it is necessary', *il semble* 'it seems', *il vaut mieux* 'it would be better', etc.), and finally, subordinate expressions labelled as '*perspectivantes*'. This last category refers to declarative verbs, opinion verbs, or verbs expressing certainty which are negated or interrogatives (e.g., *comprendre* 'to understand', *espérer* 'to hope', *expliquer* 'to explain', etc.), as well as those which denote an

inherent negative meaning (e.g., *douter* 'to doubt', *ignorer* 'to ignore', etc.). These divisions were not clearly motivated and grouped together many different subjunctive governors, potentially masking lexical effects. In fact, the exact identity of all of governors included in each of the categories or the rates associated with each was not revealed.

Lagerqvist found that the subjunctive was nearly categorical in all subgroups with the exception of the last, where negated and interrogative constructions only triggered the subjunctive at rates between 27-35%. Without a breakdown of the different types of governors within this category, it is impossible to explain this result. He also found categorical subjunctive use after the adverbial conjunctions *pour que* 'so that', *afin que* 'so that', *que* 'that', *avant que* 'before', and *jusqu'à ce que* 'until'.

He concluded that subjunctive use in subordinate clauses is abundant in written HF and argued that it is, in most of the contexts examined, conditioned by the semantic notions of futurity, irrealis, modal symmetry, and potentiality. He also suggested that the use of the future (i.e., *il faut qu'il* ***aille*** (S) 'it is necessary that he go' vs. *il faut qu'il* ***ira*** (F) 'it is necessary that he will go') is an acceptable and unsurprising subjunctive replacement strategy since both the subjunctive and the future can be used to express actions which have not yet occurred and are more or less equivalent to one another (ibid., 164-65).

Sand (2003; 1983)

Sand (2003; 1983)[27] used a small (unjustified) sample from one of the corpora which form the basis of this study (ESLO1-Orléans, 1968-70) in order to compare the use of the subjunctive amongst adults with those he obtained from his 1983 study, which focused on young students and apprentices (using the Caen Corpus, recorded by him). He provides raw numbers

[27] Most of the results from Sand (1983) mentioned here are those discussed in his 2003 paper.

for all unambiguous subjunctive forms as well as indicative and ambiguous ones in his study. But like others, he provides rates of subjunctive use under a variety of governors that are structurally and semantically categorized without providing justification for why those governors were selected. They were obviously pre-selected (and likely based on prescriptive sources) because some had no subjunctive occurrences at all and were categorically associated with the indicative (e.g., *quoique* 'whatever', *non pas que* 'not that').

Context	**Rate (% / n)**
VERBAL GOVERNORS	
Subject governors (e.g., *être* +attribute 'to be +attribute')	77% (17/22)
Object governors	
Falloir 'to be necessary/have to'	98% (61/62)
Vouloir 'to want'	100% (18/18)
Other volitional verbs (e.g., *préférer* 'to prefer', *demander* 'to ask', etc.)	100 % (35/35)
Verbs of emotion (e.g., *craindre* 'to fear', *regretter* 'to regret', etc.)	90% (9/10)
Verbs of opinion/declarative (e.g., *croire* 'to believe', *penser* 'to think', *dire* 'to say', etc.)	50% (17/34)
Double modality verbs (e.g., *comprendre* 'to understand')	31% (4/13)
TOTAL	**83% (161/194)**[28]
NON-VERBAL (adverbial) GOVERNORS	
Temporal (e.g., *jusqu'à ce que* 'until', *avant que* 'before', etc.)	88% (7/8)
Consecutive (e.g., *de façon (à ce) que* 'so that', etc.)	88% (7/8)
Causal (e.g., *c'est pas que* 'it's not that', etc.)	75% (3/4)
Hypothetical (e.g., *à moins que* 'unless', etc.)	89% (8/9)
Concessive (*e.g., malgré que* 'despite', *bien que* 'even though', etc.)	60% (3/5)
Final (e.g., *pour que* 'so that', etc.)	100% (20/20)
Other	90% (9/10)
TOTAL	**89% (57/64)**

Table 3.5: Rate of subjunctive in subordinate clauses in Sand (2003).

[28] These figures are taken from the text of the article and differ from Sand's totals in his summary chart (2003:201).

His results are divided according to the different subjunctive contexts found in the data: main clauses introduced by *que*, complement clauses (which are subcategorized according to semantic class of verb), relatives, and adverbial clauses. In the case of the main clause contexts, he noted that he only found 16 examples. Under relatives, he found a rate of 78% subjunctive (n= 14/18). The rate of subjunctive was very high after both verbal and non-verbal governors (88% and 82% respectively) (Table 3.5). Sand (2003) also examined the identity of the embedded verb under some of these categories and found that *être, avoir, faire,* and *aller* represented 63-77% of the subjunctive morphology. However, in at least one case, he conceded that the same verbs account for a similar proportion of indicative morphology since they are simply amongst the most frequent in the language, thereby nullifying the apparent effect.

He concluded that the subjunctive was used much more in the ESLO subsample (85%, n=291) than in the corpus of young speakers analyzed in 1983 (Caen= 60%, n= 391) following the same methodology. Sand attributed this difference to the fact that ESLO is made up of mostly adults (who have learned and perfected subjunctive use), while the Caen corpus is made up of young speakers (who have not yet perfected it). Regardless of the rate discrepancies, it is clear in both cases that the subjunctive was most likely to be used in specific contexts that Sand categorized semantically, but which could also be construed as lexical effects. In addition, he emphasizes that the subjunctive has not disappeared from modern French (1983:312).

Dreer (2007)

Dreer's study of subjunctive use in French is based upon the analysis of examples extracted from various plays and novels. He approached the study with the hypothesis that subjunctive use is based on the level of commitment a speaker has when referring to a particular

occurrence, as well as their level of objectivity and of assertion (ibid., 75-84) (i.e., that it is semantically motivated.

While a large portion of the study was dedicated to analyzing mood choice in individual examples, Dreer did provide quantitative results for a handful of French novels and plays where he compared mood choice in the speech of different characters and different portions of the story. For example, in one of his analyses, he predicted the part of the novel that dealt with a moral dilemma and the search for a solution would favour the subjunctive, while those that dealt with "misunderstanding, intransigence or the definitive settlement of disagreements" would favour the indicative (ibid., 195). He confirmed his hypothesis that subjunctive use is semantically motivated and that it is declining over time. He also suggested that in Old French, the subjunctive had a broader meaning and was thus used in more contexts, whereas by the 20th century, its meaning had narrowed and was being used in fewer contexts. He, like others (e.g., Bauche 1929), also suggested that in present-day French, speakers avoid potential subjunctive-selecting contexts by using infinitival constructions. But unlike other researchers who have made similar claims regarding the status of the subjunctive in contemporary French, Dreer supported his conclusions with a diachronic analysis. He compared two novels in their original versions (11th and 12th centuries) with their current version (20th century) and found a decrease in the use of both the imperfect and present subjunctive (ibid., 215). However, he only provides raw numbers of subjunctive occurrences so it is not entirely clear if they were replaced by the indicative, the conditional, infinitival, or by other variants.

3.2.2.2 Summary

This overview of several quantitative studies of the subjunctive in HF has revealed some interesting trends. First, some conclude that the subjunctive has declined in contemporary HF

(e.g., Dreer 2007; Vandergheynst 1982), while others claim that it has not (e.g., Lagerqvist 2003, Nerio 1978), yet none of the studies are based on diachronic evidence (with the exception of Dreer, whose claim is based on the comparison of two novels across time -- a time frame that is not entirely relevant to the questions being investigated here). Second, some of the studies claim that subjunctive use is at least to some degree semantically motivated (e.g., Lagerqvist 2003; Vandergheynst 1982; Nordahl 1969), while others (e.g., Sand 2003; 1983; Blanche-Benveniste et al. 1990; Gougenheim et al. 1964) use semantic and/or structural classifications to present their results but do not necessarily comment on their semantic value. Finally, these studies found widely different overall rates of subjunctive use, from around 5% (Chauveau 1998; Bieler 1972) to over 80% (e.g., Sand 2003) with many others that fall in between.

These inconsistencies stem from several different factors. One factor involves the type of data upon which the results are based, which varies from study to study. Some are based on written language (e.g., Dreer 2007; Lagerqvist 2003; Nordahl 1969) while others are based on spoken data (e.g., Sand 2003; 1983; Blanche-Benveniste et al. 1990; Vandergheynst 1982). These two sources can be expected to differ greatly from one another in terms of the rate of subjunctive within them since written language tends to adhere more closely to prescriptive norms than spoken language. In addition, many of the studies that use speech corpora as their data source fail to justify the subsamples of speakers analyzed. For example, Vandergheynst (1982) makes use of the Orléans corpus (ESLO1) but does not comment on how many speakers were selected or why. His very low token count would suggest that only a few speakers were sampled but this is never confirmed. Additionally, the Orléans corpus is comprised of speakers who lived in Orléans at the time and, therefore, includes speakers who had moved there recently from other French-speaking areas or countries. Gougenheim et al. (1964) base their analysis on a

corpus of casual interviews with French speakers who originated from all over France and even from other French-speaking countries. Chauveau (1998) uses written sources from Haute-Bretagne, while Lagerqvist (2003) makes use of Parisian newspapers.

A more serious factor that contributes to the inconsistent results across HF studies is the definition of the variable context (i.e., inconsistencies in the contexts that were included in the study). For example, Gougenheim et al. (1964) describe subjunctive use under ten verbal and non-verbal governors, while Vandergheynst (1982) analyzes its use under potential subjunctive-selecting subordinate clauses introduced by a verbal governor and *que*. Bieler (1972) uses all subordinate clauses, including non-subjunctive-selecting ones, in much of his analysis, while Chauveau (1998) limits himself to examples from lexical and regional dictionaries. In addition, there are differences in the identity of the variants included in the studies. Gougenheim et al. (1964) include subjunctive occurrences but does not differentiate between the indicative and ambiguous embedded verbs, which can lead to deflated or inflated subjunctive rates, depending how they are grouped. Vandergheynst (1982) classifies the ambiguous occurrences as either subjunctive or indicative based on which of the two appear most often in each context he analyzes, while Blanche-Benveniste et al. (1990) only report on subjunctive occurrences and make no mention of non-occurrences (i.e., the indicative or conditional).

Finally, the semantic and/or structural classification of subjunctive occurrences varies across studies. In some cases, the researcher adopted pre-established prescriptively motivated classifications and analyzed subjunctive use within them (e.g., Sand 2003; 1983), while in other studies, new categorizations were established based on the data (e.g., Lagerqvist 2003; Nordahl 1969). In addition, nearly all of the quantitative studies that suggested directly or indirectly that the subjunctive was semantically motivated (e.g., Lagerqvist 2003; Sand 2003; 1983;

Vandergheynst 1982) mainly do so on the basis of the rates of occurrences within these pre-established categories. One exception is Dreer (2007), who analyzed individual utterances according to their degree of assertion, level of objectivity, and of commitment. All of these inconsistencies across studies not only limit comparability, but also contribute to a fairly cloudy picture of how the subjunctive is actually used in contemporary HF.

Despite such divergences, some interesting findings have nonetheless emerged. In all the studies that classified the governors according to semantic class (e.g., Lagerqvist 2003; Sand 2003; 1983; Vandergheynst 1982), emotive and/or volitive verbal governors were more likely to co-occur with the subjunctive than verbs of opinion and/or perception. Also, *pour que* and *falloir* both emerged as the most frequent non-verbal and verbal governor respectively and both are associated with nearly categorical rates of subjunctive in the cases where this information is provided, although it is not usually explicitly stated. Sand (1983:308) suggested that there may be a correlation between subjunctive use and the distance between the governor and embedded verb, although he did not test this empirically. Nordahl (1969) alluded to the possibility of a similar effect on a very specific subset of his data. Sand was the only one to consider the lexical identity of the embedded verb, and found that *être, avoir, aller,* and *faire* accounted for approximately two thirds of the embedded verbs with subjunctive morphology in most contexts, although he also found a similar ratio of these verbs with indicative morphology.

The aim of this study is to provide a clear and detailed picture of how the subjunctive is used in HF at the community level, a goal that will be achieved through the application of the variationist method. The analysis will be based on vernacular speech from a carefully selected subsample of speakers who are stratified according to age and sex. Moreover, the variable context will be defined according to where the subjunctive is actually used and not where it

should be used according to prescriptive norms. Importantly, this method will allow for the simultaneous testing of the effects of both semantic and structural factors, some of which have been alluded to in the literature, on subjunctive use thereby transcending overall rates of occurrence (although rates of occurrence will prove to be very useful in the analysis). Finally, employing an accountable method will allow for a detailed comparison of subjunctive use in HF with its use in QF (already established in detail by Poplack and her colleagues), which will help to identify any (dis)similarities between the two varieties, particularly with regards to the subjunctive's productivity both in terms of its frequency of use across different contexts and the degree to which it is semantically motivated.

3.2.3 Variationist studies of the subjunctive in Canadian French

The use of the variationist framework to study subjunctive use is not new in the Canadian French context. It has been employed in several studies that have looked at different varieties of Canadian French. The most extensive analysis, undertaken by Poplack, Lealess & Dion (2013) (including data from Poplack 1992; 1990, and St-Amand 2002) looked at subjunctive use in spontaneous spoken QF across a timespan of nearly a century and a half. Despite reports of the recession of the subjunctive in contemporary speech (e.g., Laurier 1989, for Ontario French), Poplack, Lealess & Dion (2013) actually found an increase in its use across time. While the subjunctive was used variably with dozens of governors, this increase was actually due to the increasingly strong association between the subjunctive and only a few verbal governors (*falloir* 'it is necessary', *vouloir* 'to want', *aimer* 'to like'), a few non-verbal governors (*pour que* 'so that', *mais que* 'as soon as', *avant que* 'before') and a handful of embedded verbs (*être* 'to be', *avoir* 'to have', *faire'* to make', *aller* 'to go', *pouvoir* 'to be able'), despite it being prescribed with hundreds of verbal and non-verbal governors (Poplack, Lealess & Dion 2013).

Auger (1990) also studied mood alternation in subordinate clauses in spoken Quebec City French from a variationist perspective, focusing specifically on impersonal constructions. She found that some constructions co-occurred categorically with the subjunctive or with the indicative and that only a few exhibited some degree of variability. Certain impersonal constructions (e.g., *il faut que* 'it is necessary that') exhibited very high rates of subjunctive and also accounted for a large proportion of the data (>50%), which is parallel to Poplack and her colleagues' findings in Ottawa-Hull (ibid., 86). She also noted a correlation between certain semantic categories of impersonal expressions and the (non-)use of the subjunctive. For example, she found that appreciative expressions (*c'est bien/choquant/naturel* 'it's really good/shocking/natural', *ça vaut la peine* 'it's worth the trouble') tended to occur with the subjunctive, while declarative expressions (*c'est vrai/certain/dit* 'it's true/certain/said', *ça a l'air* 'it appears') and frequency expressions (*c'est rare/fréquent* 'it's rare/frequent') were more likely to occur with the indicative (ibid., 118). Auger considered various conditioning factors in her analysis, including the tense of the matrix clause, the presence of *que*, distance between the governor and embedded verb, amongst others, but found little effect on variant choice.

Laurier's (1989) variationist study of the subjunctive in Ontarian French revealed that it was the frequency of the governor that contributed the most to subjunctive selection. He also found that the linguistic dominance of the speaker had an effect on choice of mood (i.e., French-dominant speakers use more subjunctive than other speakers). However, his statistical results are marred by some severe methodological shortcomings, including some strong interactions. These are outlined in more detail in St. Amand (2002).

Grimm (2015), also employing variationist methodology, found that in Ontarian French, subjunctive use was quite robust amongst certain groups of speakers (categorized according to

level of language restriction), particularly in specific lexically-defined contexts, many of which were identical to those identified by Poplack (1992; 1990) and her colleagues (2013).

Unlike the others mentioned thus far, Comeau (2011) found no variability at all in his analysis of the subjunctive in Acadian French: the subjunctive was used nearly categorically in the contexts in which it appeared, which led him to conclude that it remains highly productive in this variety (despite the fact that it is used with only a handful of governors and is not used at all with hundreds of other governors).

Most of these studies used the community norm as the basis for subjunctive use, which means that they only considered contexts in which subjunctive morphology was actually used rather than relying on prescriptive norms. As such, they were able to provide a clear picture of its role in contemporary Canadian French speech.

In addition, the variationist method allowed for the discovery of two important findings with regards to French spoken in Canada. First, subjunctive use is variable. In other words, it is not used categorically in the contexts in which it is prescribed or in which it is found. Secondly, there is strong evidence of lexical and structural conditioning coupled with an absence of semantic conditioning.

Comeau's (2011) interpretation of his results in the variety of Acadian French he studied is an exception to both of these findings: He reported categorical use of the subjunctive in most of the contexts he analyzed and evidence of semantic conditioning in the single context in which he found variability (ibid., 177). However, Comeau did find that subjunctive use was restricted to a handful of governors, which in and of itself could be arguably construed as a lexical effect, although he did not interpret it this way. Auger (1990) found the semantic category of an impersonal expression to contribute to the likelihood of a subjunctive being used but she also

found that both lexical and structural effects were at play too, albeit minimally. Importantly, these studies have shown that subjunctive use is not in decline in Canadian French. Rather, it is strongly associated with specific lexically- and structurally-defined contexts and this association is increasing, at least in Poplack, Lealess & Dion's (2013) study of QF (the only one that includes a diachronic component).

3.2.4 This study

This study represents the first variationist study of subjunctive use in mainstream HF, as well as the first in-depth comparative analysis of its use in two varieties of French (HF and QF) across two periods of time (the 20th and 21st centuries). In addition to the rates of occurrence of the subjunctive, the internal (linguistic) and external (social) factors that could potentially condition its use will also be examined in order to identify any (dis)similarities between the two varieties. The analysis will aim to answer the following questions:

1) Is the subjunctive used to the same degree in HF as in QF across time?

2) Are there parallels in the linguistic conditioning of the variability in both varieties across time?

Answering these questions through a horizontal comparison of subjunctive use in HF and QF will help to address the larger-scale question of this dissertation: Is the variable use of the subjunctive inherited from a common source or does it result from linguistic innovation that occurred in one of the two varieties? Recall that if similarities are found in the underlying grammar of subjunctive use in the two varieties, this would suggest that the variability was inherited from a common source, or it could be indicative of a shared innovation. Differences

could suggest that one variety underwent a linguistic innovation, or that innovation occurred in both but proceeded along different paths. These ideas will be explored in the ensuing sections.

3.3 Methodology

3.3.1 Circumscribing the variable context

Any variationist analysis centres on the construct of the linguistic variable, which refers to alternate ways of "saying the same thing" without a change in meaning (Labov 1982:22; 1966). The choice of verbal mood does not appear to fit this definition since many ascribe various meanings to the choice of the subjunctive over the indicative or the conditional. While it is known that the subjunctive can only occur in specific contexts in contemporary French (namely in subordinate clauses headed by certain governors), the exact nature of these constructions remains unclear since they tend to vary both in prescriptive and descriptive literature. Following Poplack and her colleagues, the variable context is defined here by identifying the contexts in which it is actually used, not where it is prescribed, and by noting all of the other tenses that occur in these contexts (i.e., variants). By defining the variable context in such a way as to include all of the contexts where the subjunctive can be used, it is then possible to test whether or not the use of one variant over another does, in fact, stem from a change in meaning. It will also help to establish, independently of prescriptive dictates, how the subjunctive is used by the HF community.

The first step in circumscribing the variable context involved locating all of the unambiguous subjunctive forms (as in (7)) in the data and taking note of each of the governors, both verbal and non-verbal, that triggered at least one unambiguous subjunctive form in the embedded clause. The variable context is limited to subjunctives embedded under verbal or non-

verbal governors in this study for comparative purposes, but, so as to not miss any major trends, main clause (8) and relative subjunctives (9) are examined separately in Section 3.4.1.4.

(7) « J'pense pas qu'ils l'***aient*** (S) fait cette année. » (CFPP/GP/732)

'I don't think they've done it this year.'

(8) « Non parce que s'il veut aller faire du bateau qu'il ***sorte*** (S) d'Orléans et qu'il ***aille*** (S) à Olivet. » (ESLO1/1268/1207)

'No, because if he wants to go boating, he can leave Orléans and go to Olivet.'

(9) « Je pense pas que la façon dont ils ***soient*** (S) éduqués euh soit[29] (S) à renier à cent pour cent. » (ESLO2/UZ57/528)

'I don't think that the way in which they were educated should be one-hundred percent disregarded.'

Next, all occurrences of each of the governors (e.g., *penser* 'to think' in (7)) were located and all the other embedded forms, including conditional (past and present) and indicative (present, *passé composé, imparfait*, *plus-que-parfait*, future, etc.) embedded under them (10, 11) were noted. These are the *non-applications* (i.e., variants). In this way, the *Principle of Accountability*,[30] which requires that one considers not only contexts in which the subjunctive occurs but also those in which it could have even if it does not (Labov 1972), is respected.

As shown in examples (7) to (11), the morphological form of the embedded verb is variable in certain subjunctive-selecting contexts. But is this variability determined by a change

[29] This subjunctive occurrence is included as embedded under the verbal governor *penser* 'to think'.

[30] I am aware of the caveat that potential governors that only occurred with indicative and/or conditional tokens are not represented in the analysis, but the purpose of this study is to focus on contexts in which the subjunctive is actually used, not where it is prescribed.

If the subjunctive occurrence was found under a negated governor (e.g *penser*) only non-applications that occurred under the negated governor were considered. The same is true for affirmative governors (e.g., *souhaiter* 'to wish'). If the subjunctive occurred in both negative and affirmative cases (e.g., *falloir*), non-applications were extracted for both affirmative and negative contexts. This was corpus-specific.

in meaning? Or are other factors at play? These are some of the questions that this study seeks to answer.

(10) « Je pense pas que ça ***a*** (I) beaucoup changé. » (ESLO1/QB1000/477)

'I don't think that it has changed much.'

(11) « Non, je ne pense pas je ***ferais*** (C) de brouillon. » (ESLO1/IZ248/1376)

'No, I don't think that I would make a rough copy.'

In sum, the envelope of variation includes any context in which the subjunctive could have appeared (whether it did or not) under any verbal or non-verbal governor that triggered at least one subjunctive occurrence (at a rate of 2% or more) in each of the corpora[31],[32],[33]. Each occurrence is referred to as a *token*.

3.3.2 Exclusions

All ambiguous forms (including all first group *-er* verbs) were excluded from the variable context since their subjunctive forms are not distinguishable from the indicative forms, as is illustrated in Section 3.1. As a result, more than one third of the data had to be discarded from this study. This is crucial because including ambiguous tokens and counting them as subjunctives would artificially increase the overall rate, whereas if they were counted as non-applications, it would have the opposite effect. In neither case could an accurate assessment of the overall rate of subjunctive usage be obtained. These figures, presented in Table 3.6, are slightly less than the

[31] The extraction process was applied to each corpus separately. In other words, if a particular governor was found to trigger a subjunctive in one corpus only, the non-applications were extracted from this corpus and not from the others.
[32] In order to be included in the variable context, the rate of subjunctive has to be 2% or greater under any governor in order to avoid potential skewing of the results.
[33] Non-standard subjunctive forms were included as long as they were unambiguous (e.g., *alle* 'go').

50% of ambiguous form-based exclusions noted by Poplack, Lealess & Dion (2013:163) in QF but nonetheless represent a large portion of the data that had to be discarded.

	Unambiguous tokens	**Ambiguous tokens**	**% Ambiguous tokens**
Verbal governors	1155	721	38%
Non-verbal governors	292	202	41%

Table 3.6: Distribution of ambiguous and unambiguous forms in HF corpora.

In addition to ambiguous embedded forms, several other contexts were also excluded from the variable context. Fixed expressions (12), answers to questions (13)[34], as well as ambiguous or otherwise doubtful tokens[35] were excluded from the quantitative portions of the analysis because it is impossible to ascertain whether it is a true subjunctive occurrence.

(12) « J'ai pas ni l'opportunité ni quoi que ce ***soit*** (S) qui m'attire à à aller vers là-bas. » (CFPP/LD/647)

'I neither have the opportunity nor anything else that would entice me to go there.'

(13) « [Int] Qu'est-ce que vous attendez de lui ? [098] Ben qu'il nous ***défende-*** (S) qu' il ***défende*** (S) tous nos droits. » (ESLO1/SD272/566)

'[Int] What do you expect of him? [098] Well that he defend us- that he defends all of our rights.'

This process resulted in a total of 1447 tokens of subjunctive-selecting contexts hosting at least one unambiguous subjunctive (at a rate of at least 2%) that were retained for analysis.

[34] Answers to questions were excluded not because there is any question about the use of an unambiguous subjunctive, but because the governor and the embedded clause were uttered by two different people.

[35] Certain verbs, when combined with 2nd person morphology, are often used as discourse markers in running speech (e.g., *vous voyez, vous comprenez, tu vois, vous savez, tu sais*, etc.). These were excluded from the analysis. Also, certain non-verbal governors (e.g., *après que, avant que, puis que, bien que, maintenant que, quoi que, même que, fait que*) were only extracted if the complementizer *que* was present because without it, the resulting ambiguity prevented the inclusion of the token.

3.3.3 Factors under analysis

The data were coded for various internal (linguistic) and external (social) factors in order to ascertain, through a multivariate analysis, which of these best account for variant choice. A multivariate analysis considers all factors simultaneously, and thus has the capacity to reveal whether subjunctive selection in HF is best predicted by semantic or structural factors, as well as the extent to which it differs from QF and other characterizations of HF. Internal conditioning (i.e., the portion of the grammar underlying variant choice) is a crucial component in all variationist studies of the subjunctive. The factor groups tested here are inspired by Poplack, Lealess & Dion's (2013) study of QF and include the following: semantic class of the governor, sentence type, the presence of other indicators of non-factual modality, the reality of the predication, the tense of the verbal governor, the morphological form of the embedded verb, presence of complementizer *que*, as well as the distance between the governor and the embedded verb. Each is explained below, along with supporting examples taken directly from the corpora. In order to identify possible change in progress, each token was also coded for the following external (social) factors: sex, age, and time period.

3.3.3.1 External (social) factors

As was reported earlier, there are differing accounts as to the status of the subjunctive in contemporary French, particularly as to whether it is declining or increasing. In either instance, if there is change underway, the analysis of certain external factors may allow it to be identified. The three external factors that will be tested in the context of this study are time period, sex, and age. Comparing subjunctive use across two time periods (the 20th and 21st centuries) may allow for the identification of change in progress in real time, while comparing the different groups

from the same time period could reveal change in apparent time (Table 3.7). Evidence of a decline in subjunctive use would be produced if the 21st century speakers use less of it than 20th century speakers, while with an increase, the opposite would be expected. Similarly, if young speakers use fewer subjunctives than older speakers, this would suggest that subjunctive use may be declining, while if they use more, this may indicate that it is on the rise.

	20th (Orléans)	**21st (Orléans Paris)**
15-34	26	21
35-54	17	17
55+	20	28
TOTAL	**63**	**66**

Table 3.7: Age groups and time periods considered in the present study.

The sex of the speaker was also considered in the present study (Table 3.8), due to the fact that in the case of change in progress, women tend to be at the forefront. Therefore, if clear differences across these two cohorts are found, this could be indicative of change in progress.

Code	**Sex**	**# speakers 20th C**	**# speakers 21st C**	**Total**
M	Male	32	27	59
F	Female	31	39	70
	Total	63	66	129

Table 3.8: Speaker sex categories considered in the present study.

3.3.3.2 Internal Factors (Semantic)

Both semantic and structural influences on subjunctive use are tested in this study. The semantic factor groups were included in order to address widely accepted claims that the subjunctive is triggered by specific semantic contexts. Since it is impossible for an analyst to determine the intentions or internal motivations of a speaker, the only way to test this hypothesis is to operationalize the semantic nuances that are commonly associated with the subjunctive

(e.g., doubt, irrealis, uncertainty, etc.). Poplack, Lealess & Dion (2013) developed three such factor groups designed to do just this, which are independent of the morphology found on the embedded verb: sentence type, other indicators of non-factual modality, and reality of the predication, which are replicated here. A clear semantic effect (i.e., that the subjunctive occurs freely without restrictions in specific, semantically-defined contexts) would support the idea that subjunctive use is productive.

Sentence Type

One way to operationalize the semantic hypothesis is through the degree of assertion of the main clause, which can be expressed by the sentence type. Each verbal and non-verbal token was coded according to whether the main clause was affirmative, negative, conditional, or interrogative. Interrogative, conditional, and negative sentences often encode a [-assertive] reading (Quirk et al. 1985) and, therefore, would be most likely to trigger a subjunctive. An example for each of these sentence types is given in Table 3.9.

Sentence type	Example
Affirmative	« J'sais pas, il faut que je ***réfléchisse*** (S) à la question. » (CFPP/GC/2554) 'I don't know. I should think about the question.'
Negative	« Faut pas ça ***va*** (I) pas si vite que ça. » (ESLO2/OS6/1109) 'It can't go *that* fast.'
Conditional	« Faudrait que je me ***remette*** (S) aux études. » (ESLO1/NS571/52) 'I would have to return to my studies.'
Interrogative	« Qu'est qu'il faudrait qu'ils fassent quoi pour euh réussir? » (ESLO1/SL312/327) 'What would they have to do to uh succeed?'

Table 3.9: Sentence type categories considered in the present study.

Hypothesis: According to the commonly cited semantic hypothesis (e.g., Hooper 1975) *subjunctive use is associated with non-assertion, therefore, one can expect to find more subjunctive in conditional, interrogative, and negative contexts than in affirmative ones.*

Other indicators of non-factual modality

In addition to sentence type, verbal governor tokens were coded for the presence of other indicators of non-factual modality. These indicators include specific lexical items (e.g., *peut-être*), specific tense selection (e.g., conditional), as well as the use of modal auxiliaries (e.g., *devoir*) in the utterance (Table 3.10).

Indicators of non-factual modality	**Example**
Lexical (*possiblement, probablement, si-*conditional, lexical interrogative *est-ce que*, *penser* in previous clause, *je sais pas, je suis pas sûr/e, certain/e*	« Moi je pense pas qu'on qu'on ***ait*** (S) un langage spécifique. » (ESLO2/PT45/892) 'Me, I don't think we we have a specific language.' « Je sais pas moi, il me semble qu'à âge égal maintenant on ***apprend*** (I) beaucoup plus. » (ESLO1/DT78/371) 'I don't know, it seems that at the same age today we learn a lot more.'
Tense (subjunctive, conditional, future, imperative)	« Il aurait fallu qu'elle ***prenne*** (S) des cours ouais. » (ESLO2/OS6/986) 'She would have had to take courses, yeah.'
Modal (*devoir, pouvoir*)	« J'peux comprendre que vous ***fassiez*** (S) le l'amalgame. » (CFPP/MC/417) 'I understand that you would mix the- that up.'
More than one	« Je pense que qu'il faudrait qu'il y ***ait*** (S) des changements très importants. » (ESLO1/KE383/379) 'I think that- that there would have to be some very big changes.'
None	« Il faut qu' tu ***fais*** (I) ci faut qu' tu ***fasses*** (S) ça. » (CFPP/AH/4632) 'You have to do this you should do that.'

Table 3.10: Indicators of non-factual modality under verbal governors considered in the present study.

Hypothesis: If subjunctive use is conditioned by semantic factors, it follows that if one or more of these indicators of non-factual modality are present in the sentence, there is a greater likelihood that the subjunctive will be used.

Realis/Irrealis

Non-verbal governors were also coded according to whether the embedded clause was either irrealis or realis in order to contrast events that had occurred with those that were hypothetical or had simply not yet occurred. If the event occurred before speech time, it was coded as realis, whereas if it occurred after speech time, it was coded as irrealis. If the event in the embedded clause took place or was likely to take place after the time expressed in the main clause, it was also coded as irrealis[36]. An example of each is given in Table 3.11.

Reality of predication	**Example**
Realis	« Malgré qu'c'***est*** (I) à Saint-Ouen, ça appartient à Clignancourt. » (CFPP/AH/2323) 'Even though it's in Saint-Ouen, it belongs to Clignancourt.'
Irrealis	« Je bats, je mets du lait pour que ça ***puisse*** (S) gonfler. » (ESLO2/MF363/1566) 'I beat it, I add some milk so that it can become fluffy.'

Table 3.11: Reality of the prediction categories applied to non-verbal governors considered in the present study.

Hypothesis: In the same vein as the other semantic factor groups, if there is a semantic effect on subjunctive use, higher rates of subjunctive can be expected in irrealis contexts as these refer to events that have not yet occurred and that are, by definition, non-factual.

[36] This same distinction was accounted for with verbal governors through verbal tense. Verbal tense is classified as a structural factor group (and not a semantic one) in this study for reasons explained below.

3.3.3.3 Internal Factors (Morphosyntactic)

The other factor groups are purely morphosyntactic in nature and are designed to test the influence of the surrounding structure on variant choice. If subjunctive use is best predicted by structural or lexical factors, this would suggest that it does not occur freely, without restriction, in all eligible contexts, and thus it can be inferred that its productivity is limited. In order to test whether or not this is the case, various linguistic factors are considered including the tense of the verbal governor, the nature of the embedded verb, the presence of the complementizer *que* and/or intervening material. Each factor group is explained in detail below. In order to ensure comparability, the internal factors considered in the present study are the same as those operationalized and analyzed in Poplack, Lealess & Dion (2013).

Tense of the verbal governor

Each token was coded according to the tense of the verbal governor in order to determine whether tense concordance has any effect on subjunctive use[37]. Tense concordance refers to the matching of the tense of the embedded verb to that of the matrix verb, an effect that can only be found under verbal matrices. Prescriptive rules stipulate that in French, the embedded subjunctive verb must agree with the matrix verb but this rule is usually limited to past tense contexts and is rarely found outside of literary works since the imperfect and pluperfect subjunctive have virtually disappeared from QF and HF spoken varieties of French (e.g., Glatigny 1976). Therefore, it is expected that the present (or composed: e.g., *ait eu)* subjunctive will appear under a variety of verbal matrix tenses.

[37] Verbal tense could be classified as either a structural or semantic factor group. It is included under the structural factor groups here based on the results of Poplack, Lealess & Dion (2013), who found a strong tense concordance effect to be operative in QF.

Poplack, Lealess & Dion (2013) found the subjunctive to be strongly disfavoured with conditional matrices in 19th and the 20th century QF. First noted in Poplack (1990), this was due to the presence of a particular tense concordance effect, which deviated from the standard tense concordance rule described above: conditional matrices were most likely to be followed by an embedded conditional verb, which she refers to as a non-standard "remnant of the Classical French period" (Poplack 1992:248). An example is given in (14).

(14) « ***J'aimerais*** (C) ça que le bon Dieu ***viendrait*** (C) me chercher. » (040/2559)

'I'd like it if the good Lord would come to get me.' (cited from Poplack 1992: 254)

She also noted that this pattern extended to other verbal matrices as well: the tense of the embedded (non-subjunctive) verb matched that of the matrix more often than not. Examples of each of the different verbal governor tenses that were coded in this study are given in Table 3.12.

Tense of matrix clause	**Example**
Present indicative	« Faut y ***a*** (I) quand même des priorités quoi.» (ESLO2/HF8/567) 'There still has to be priorities you know.'
Imparfait	« On a dû lui expliquer pourquoi il fallait qu'elle ***fasse*** (S) la grève. » (ESLO1/FA192/410) 'We had to explain to her why she had to go on strike.'
Passé composé	« Il *a* fallu que j'en ***prenne*** (S) assez alors j'ai trouvé ces perles poires. » (ESLO1/JI306/343) 'I had to buy enough so I found these pear-shaped pearls'
Future	« Non mais faudra que je vous ***dise*** (S) pourquoi je l'ai fait. » (ESLO2/YQ52/736) 'No but I will have to tell you why I did it.'
Conditional (past or present)	« Faudrait que je me ***remette*** (S) aux études. » (ESLO1/NS571/52) 'I should get back to my studies.'

Table 3.12: Examples of verbal governor tense.

Hypothesis: If a QF-type tense concordance pattern is operative in HF, low rates of subjunctive after conditional matrices can be expected.

Morphological form of embedded verb

Each token (with verbal and non-verbal governors) was coded according to whether the embedded verb was suppletive (i.e., has an irregular subjunctive form) or regular (i.e., the subjunctive form comes from the indicative paradigm). An example of each is given in Table 3.13. Verbs coded as suppletive in the present study include *aller, avoir, être, devoir, faire, falloir, pouvoir, savoir, valoir*, and *vouloir*. Poplack, Lealess & Dion (2013) found that certain frequent suppletive verbs were generally more likely to co-occur with subjunctive morphology than regular verbs in QF. Bybee & Thompson (1997) suggest that this is due to the Conserving Effect: frequent (and in this case, irregular) lexical items often resist change and become subject to entrenchment.

Morphological form of embedded verb	**Example**
Suppletive	« Je bats, je mets du lait pour que ça ***puisse*** (S) gonfler. » (ESLO2/MF363/156) 'I beat it, I add some milk so that it can get fluffy.'
Regular	« Ils reviennent me voir pour que je les ***remette*** (S) plus en état leurs leurs outils.» (ESLO2/RW27/355) 'They come back to see me so that I can refurbish their- their tools.'

Table 3.13: Examples of the morphological forms of the embedded verb considered in this study.

Hypothesis: According to Bybee & Thompson's (1997) Conserving Effect, frequent suppletive verbs are most likely to be marked with subjunctive morphology and will resist the general pattern of decline of the subjunctive (if present).

Presence of complementizer '*que*'

The presence or absence of the complementizer *que* 'that' was also coded for each token of both verbal and non-verbal governors[38]. The complementizer *que* is considered to be a stereotypical feature of the subjunctive but is frequently deleted in vernacular speech for independent reasons (e.g., Martineau 1985). This factor group was included in order to determine the effect of its presence on subjunctive use. Examples are given in Table 3.14. Poplack, Lealess & Dion (2013) found that presence of *que* was favourable to subjunctive use in most of the cohorts examined.

Presence of *que*	**Example**
Deleted	« Je pense pas Ø y ***a*** (I) tellement de choses à apprendre maintenant.» (ESLO1/FC716/183) 'I don't think there are so many things to learn now'
Overt	« Mais de toute façon il faut qu'il ***ait*** (S) la majorité des voix.» (ESLO1/RV252/583) 'In any case, he has to have the majority of the votes.'

Table 3.14: Examples of the overt and deleted complementizer *que.*

Hypothesis: Since 'que' is considered to be a prototypical feature of the contexts requiring the subjunctive, subjunctive can be expected to be higher when 'que' is present.

Intervening material

Finally, each token (both verbal and non-verbal) was also coded for the presence of intervening material (Table 3.15)[39]. Poplack, Lealess & Dion (2013) found that the subjunctive

[38] These were extracted by first identifying unambiguous subjunctive morphology, then the subjunctive governor, and then a corresponding embedded clause. Due to the fact that these contexts were not always clear (i.e., a restart could be disguised as an embedded clause), I checked each of these tokens manually in the original recordings (where possible).

[39] The coding protocol adopted here does not consider the *amount* of intervening material, only the *type* of intervening material. One evaluator pointed out that the amount of intervening material may also play a role in the (non-)application of the subjunctive. While this may be the case, I opted to replicate the original coding protocol

was most likely to occur in contexts in which there was no intervening material between the governor and the embedded verb. Sand (1983) also noted this possibility but did not test it quantitatively across all possible contexts, while Nordahl (1969) found it to only be relevant to verbal governors he classified as subjective.

Intervening material	**Example**
Syntactic (intervening clause with clear syntactic link including other negators *ne, plus, jamais*, etc., double subject or object clitics, *ça*, etc.)	« J'avais une tante qui était là et qui voulait absolument que je lui ***dise*** (S) quelque chose. » (ESLO1/JI306/1772) 'I had an aunt who was there and who absolutely wanted me to tell her something.'
	« Il faut surtout pas que ça ***devienne*** (S) une ville musée. » (CFPP/PB1/2553) 'Above all, it shouldn't turn into a museum city'
Parenthetical (any intervening clause with no syntactic link including false starts, tags, discourse markers, conjoined clauses, and non-linked pronouns, etc.).	« C'est que tous les gens voudraient que leur que leur gosse ***soit*** (S) euh toubib ou avocat. » (ESLO1/HU339/119) 'It's that everyone would like their kid to be a uh- doctor or a lawyer.'
	« Il faut qu'il y ait quand même quel- condi- quelques conditions c'est-à-dire que j'***sois*** (S) pas très en retard. » (CFPP/MR1/1060) 'There still should be som- condi- some conditions, like that I can't be really late.'
None (includes *pas, là,* and *y* in *il y a.*)	« Il faudrait que la couturière ***ait*** (S) un statut spécial. » (ESLO1/JI306/540) 'The seamstress would have to have a special status.'
	« La première fois il faut qu'*y* **ait** (S) une majorité absolue. » (ESLO1/1254/632) 'The first time, there has to be an absolute majority.'

Table 3.15: Examples of different types of intervening material considered in the present study.

Hypothesis: Extra material inserted between the governor and the embedded verb could potentially obscure the relationship between the two and thus lead to lower rates of subjunctive.

used in Poplack, Lealess & Dion (2013). It should be noted, however, that in the majority of cases, syntactic intervening material included only one or two words (e.g., a pronoun, a negator), and parenthetical intervening material often included more than two words, sometimes entire clauses.

In the following sections, the data are analyzed according to the rates of use and the factors outlined above in order to determine which, if any, condition subjunctive use in contemporary HF.

3.4 Results

In this section, subjunctive use in contemporary 20th century HF is examined. Next, these findings are compared to those of Poplack, Lealess & Dion (2013) for QF in the same period in order to identify cross-variety differences or similarities. This is followed by a diachronic analysis intended to determine if any changes took place between the late 20th century and the early 21st century in HF and how this compares to the diachronic trajectory of subjunctive use in QF already established by Poplack and her colleagues for roughly the same period of time[40]

It is important to reiterate here, before presenting the results, that the methodological approach employed in this analysis is a replication of the one used by Poplack, Lealess & Dion (2013) in their detailed diachronic analysis of subjunctive use in QF across three time periods. One of the main findings to emerge from this study was the entrenchment of subjunctive use across time in very specific structural contexts (i.e., a lack of productivity).

This finding was partially revealed through their use of three innovative measures of subjunctive productivity, which included not only the association of the subjunctive with each governor (rate), but also the proportion that each governor represents of the entire pool of governors (% pool) as well as the proportion of all of the subjunctive morphology (% subj.) accounted for by each governor. High scores on each of these measures for specific governors would support the idea that the subjunctive is lacking in productivity. If the subjunctive is indeed productive, it should occur freely "across all lexical items and contexts forming its domain"

[40] They also include 19th century QF as a point of comparison.

(Poplack 2001:411). In other words, disparate rates coupled with disproportionate occurrences of the different lexical governors would suggest that subjunctive use has been routinized in specific contexts and thus, is not productive. For example, if one governor is very frequent (i.e., occupies a large portion of the governor pool) and also happens to be strongly associated with the subjunctive, it would follow that this same governor would also account for a large portion of the subjunctive morphology and, therefore, its role in relation to the other governors would be disproportionate.

Another measure employed by Poplack and her colleagues to test productivity was that of semanticity (i.e., whether or not the subjunctive actually expresses the meanings commonly attributed to it) (Poplack 2001:411). They tested the degree to which subjunctive use is semantically conditioned through the operationalization of several semantic nuances it is most often claimed to express (these were outlined above in the description of the factor groups included in this study). According to Poplack, if the use of the subjunctive is semantically motivated, it should occur freely everywhere the nuances are expressed (in a subjunctive-eligible context) without restriction (Poplack 2001:411). Varying rates across contexts and/or governors expressing these meanings would suggest that forces other than semantic ones are at play. If subjunctive use is best accounted for by the linguistic context (i.e., lexical and/or structural factors) and not by semantic factors, decreased productivity can be inferred.

An initial and somewhat surprising finding from their study was an increase in subjunctive use under verbal governors across three time periods in QF, with the overall rate of subjunctive rising from 59% in the 19th century to 76% in the 20th century and finally to 79% in the 21st century. However, their use of the productivity measures outline above revealed that what appeared to be a surge in subjunctive use was actually due to the increased association of

the subjunctive with one single verbal governor: *falloir*. Without it, the rate of subjunctive remained stable across time, ranging from 51%-55% in each of the three time periods examined.

Poplack and her colleagues further discovered that by the 21st century, subjunctive use was largely restricted to only a few verbal governors (*falloir, aimer, vouloir*) as well as to a handful of embedded verbs (*être, avoir, faire, aller*) and that its entrenchment in these contexts grew stronger over time (with the governor in particular). When the most favourable verbal governors were combined with the embedded verbs, subjunctive rates rose to 99% in the 21st QF data. They found a similar result under non-verbal governors in that only a few (*pour que, mais que, avant que,* and *jusqu'à ce que*) were not only very likely to co-occur with a subjunctive embedded verb but also that when combined, they accounted for the vast majority of all the subjunctive morphology found in the data set. When one of these top non-verbal governors occurred with a top embedded verb, subjunctive rates were predictably near-categorical. Thus the rates of subjunctive usage were intimately tied to and only explicable by disproportions in governors as well as in their associations with the subjunctive.

Additionally, through the operationalization of potential semantic conditioning factors, they found that both in the case of verbal and non-verbal governors, the use of the subjunctive was best predicted by lexical and structural factors and not by semantic ones, thereby reaffirming its lack of productivity in QF.

Are the patterns of subjunctive use identified by Poplack, Lealess & Dion (2013) idiosyncratic to QF or are they also present in HF? Thus far, they have not been tested elsewhere, but as was outlined in Section 3.2, most studies of the subjunctive have not employed a methodology that would lead to their identification. One of the main objectives of this study is to

test, through a synchronic comparison with HF, whether or not these patterns are, in fact, unique to QF.

3.4.1 Synchronic comparative analysis of subjunctive use in HF and QF

In this section, subjunctive use in 20th century HF is compared with 20th century QF[41]. The overall rate of subjunctive under both verbal and non-verbal governors combined in 20th century HF is 71% (Table 3.16). This means that the subjunctive is only used in less than three-quarters of eligible contexts and that it is variable, which is contrary to prescriptive norms that stipulate that the subjunctive be used in all 'eligible contexts'[42].

HF	QF
71% (n= 626/877)	76% (n= 2116/2767)

Table 3.16: Overall rate of subjunctive under verbal and non-verbal governors in 20th century HF and QF[43].

When compared to QF, the overall rate of subjunctive use in subordinate clauses under verbal and non-verbal governors combined is *lower* in HF (71%) than in QF (76%), and this difference is statistically significant, according to Fisher's exact test (p= 0.0026). This result is completely contrary to expectations since HF is often touted as hewing more closely to prescriptive dictates than QF (i.e., one would expect to find a higher rate of subjunctive in HF than in QF). This result does not closely coincide with the HF literature, where overall rates as

[41] All results reported here for QF are adapted from Poplack, Lealess & Dion (2013).
[42] Recall that the group of 'eligible' subjunctive contexts is nearly impossible to identify since they tend to vary from one grammarian to the next across time and even within the same time period (Poplack, Lealess & Dion 2013). This is why a usage-based approach was adopted here in order to identify which contexts are eligible for subjunctive use.
[43] Results for QF adapted from Poplack, Lealess & Dion (2013:165)

high as 84% (Sand 2003) and as low as 5% (e.g., Chauveau 1998; Bieler 1972) have been noted[44].

Subjunctive use under verbal and non-verbal governors will now be examined individually in order to determine to what degree these varieties differ in each context both in terms of overall rate of occurrence and the underlying conditioning factors.

3.4.1.1 Verbal governors

Overall rates

The rate of subjunctive in HF under verbal governors is 71% (Table 3.17). This rate is not as high as what was found in QF (76%) and the difference between the two is statistically significant (p= 0.0068) according to Fisher's exact test.

HF	QF
71% (n= 520/732)	76% (n= 1953/2569)

Table 3.17: Overall rate of subjunctive under verbal governors in 20th century HF and QF[45].

Is it the case that all verbal governors behave similarly to one another? Or do some exhibit proportional irregularities in their contribution to subjunctive use? The lexical identity of the verbal governors is examined in detail in the next section in order to address these queries.

Lexical identity of verbal governors

Subjunctive morphology was found under a total of 49 verbal governors in 20th century HF, which is a far cry from the handful described in Gougenheim et al. (1964) and from the hundreds Poplack, Lealess & Dion (2013) found to be prescribed over the course of several

[44] Sand's data comes from one of the corpora used here but his methodology differs from ours, hence the differences in rate of occurrence. This is detailed in Section 3.2.
[45] Results for QF adapted from Poplack, Lealess & Dion (2013) and from a personal communication.

centuries. This number is also substantially higher than the number found by Poplack, Lealess & Dion (2013) in QF (n= 34). This is particularly interesting given that the number of words in the 20th century HF subsample (approx. 800, 000) represents only a small proportion of the number of words in the 20th century QF corpus (approx. 2,500,000). One could argue that the more words in the corpus, the more likely one is to find different subjunctive governors. The frequency of subjunctive verbal governors per thousand words is much more elevated in HF (5.9) than in QF (1.3), which confirms that subjunctive morphology is spread across many more verbally-governed contexts in the former, thus hinting at a higher degree of productivity. However, the overall rate of subjunctive is actually lower in HF (71%), which would appear to contradict the suggestion of increased productivity relative to QF. This will be explored further throughout the chapter.

Recall that one of the ways productivity of the subjunctive can be established is through its substantial occurrence across all eligible contexts (Poplack 2001). This does not appear to be the case in HF: not all governors (i.e., contexts) exhibit similar rates of subjunctive (Figure 3.1). In fact, the rate ranges from 6% to 100% subjunctive, depending on the verbal governor. Interestingly, the majority of these governors co-occur (near) categorically with the subjunctive so it may seem surprising that the overall rate under verbal governors is only 71% in HF. How can this be explained?

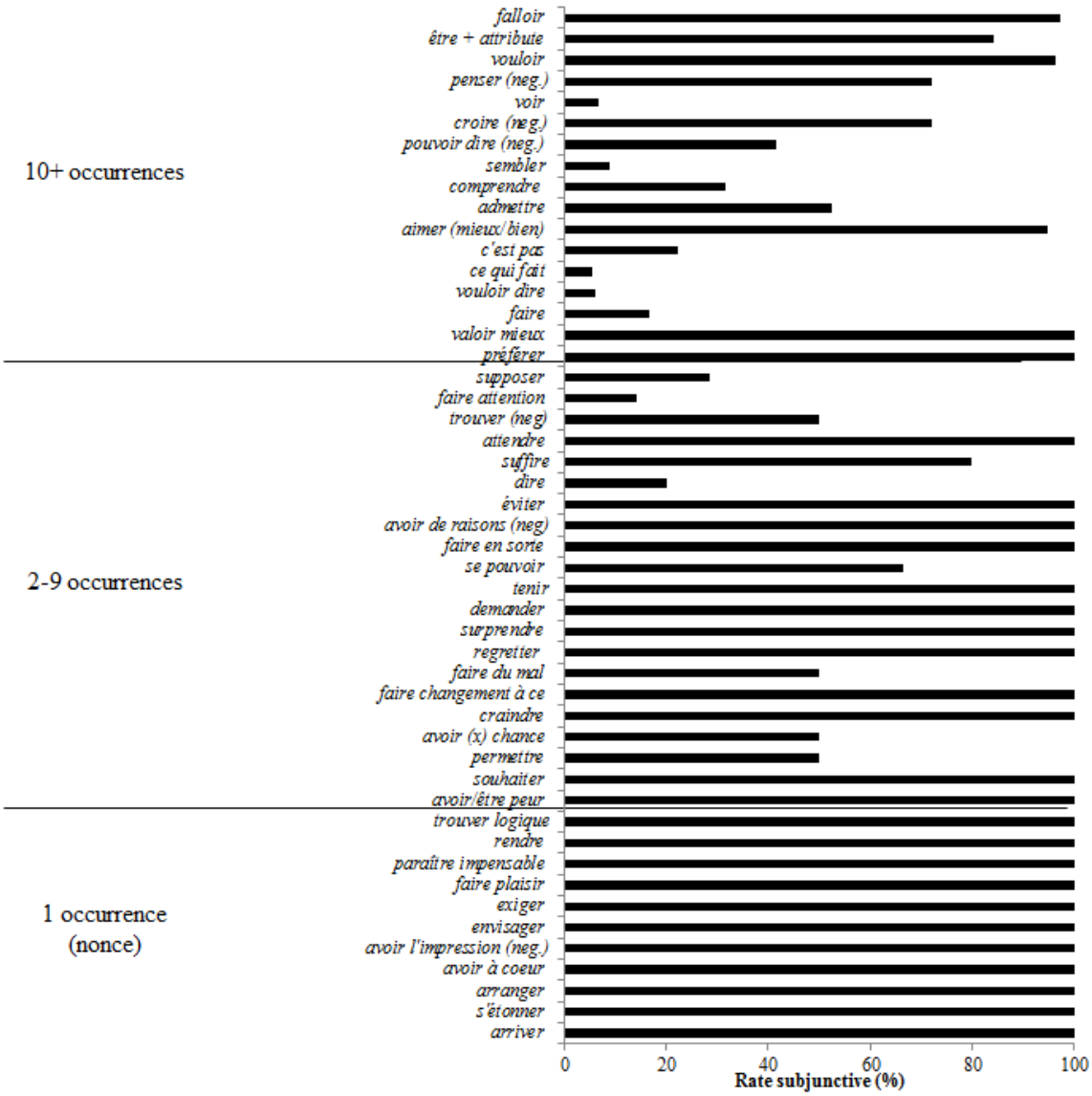


Figure 3.1: Rate of subjunctive under each individual verbal governor in 20th century HF according to overall number of occurrences of the governor (listed in descending order of frequency)[46].

A first point to make is that there are clear disproportions with regards to the rates under each governor and the frequency of the governors themselves. The display in Figure 3.1

[46] Here, and in all subsequent tables and text, when a governor is followed by (neg.), it refers to the fact that the subjunctive occurred exclusively under the negated form of this governor, which means that non-applictions were only extracted from negated contexts. With all other governors, the subjunctive either occurred in affirmative contexts only, or in both negative and affirmative ones, which formed the basis for the extraction of non-applications in those cases. A complete list can be found in Appendix A.

illustrates the rate of subjunctive according to the frequency of occurrence of each verbal governor, differentiating between high frequency verbal governors (10+ occurrences), medium frequency (2-9 occurrences) ones, as well as those which occur one single time (i.e., nonce). It is clear from these results that not all verbal governors contribute to subjunctive use in the same capacity: many of those which co-occur categorically with the subjunctive are low to medium frequency. In fact, nearly half of these (n= 11/24) are nonce governors, occurring only a single time in the data set. These facts have not been previously discussed in quantitative studies of HF.

How does this compare to what was found in QF? Poplack noted the same type of pattern in 20th century QF: different lexical governors vary wildly in terms of their overall association with the subjunctive and their proportional value vis-à-vis the other governors. The list of verbal governors in both HF and QF is given in Table 3.18. There are some other striking similarities to note across the two varieties. First, only a small proportion of the verbal governors are classified as frequent (i.e., 10 or more occurrences) in each variety (HF: 35%, QF: 38%). Second, while it is true that many of the verbal governors vary across the two varieties, there is still quite a bit of overlap, with a total of 19 shared governors. Nearly all of the frequent governors found in QF (n= 10/13) are also found in HF (though not always with the same frequency of occurrence). Third, rates of occurrence of the subjunctive vary independently of the frequency of the governor (with the exception of nonce governors, which necessarily exhibit rates of 100%). It follows that certain high frequency governors with strong associations with the subjunctive will have a stronger influence on the overall rate of subjunctive, which is discussed in more detail below.

Despite some fluctuation with respect to lexical identity, both varieties appear to share the same general pattern: only a small proportion of the verbal governors that co-occur with the subjunctive do so regularly (i.e., high rates of subjunctive under frequent governors). As

mentioned earlier, Bybee & Thompson (1997:382) suggest that the entrenchment of the subjunctive in QF with only a few frequent governors and embedded verbs is the result of the Conserving Effect, which refers to the fact that high frequency lexical items often resist change (toward the indicative in this case) and become subject to entrenchment. In an attempt to establish whether or not this is also the case in HF, some of the highly frequent and high subjunctive governors are examined in more detail below[47].

[47] A few notes to make about the HF verbal governors listed in Table 3.18:
- Data from QF originate from a personal communication from Dr. Shana Poplack. The translation for each of the verbal governors found in HF and QF is given in Appendix A.
- *Sembler '*it seems' includes occurrences with (e.g., il *me semble*) and without (e.g., *il semble*) an object pronoun, the latter being the construction that is normally prescribed with the subjunctive. A subjunctive rate of 100% (n= 1/1) was found in the simple construction (no object pronoun) and a rate of 5% (n= 1/21) in contexts with the object pronoun. Both occurrences of the subjunctive with *sembler* are given below in (a) and (b).

a. « Il semble que le français ***soit*** (S) adopté d'une façon extrêmement moderne. » (ESLO1/1268/1022)
'It seems that French is adopted in an extremely modern way.'
b. « Ca leur semble très impossible qu'une femme ne ***soit*** (S) pas là. » (ESLO1/FA192/108).
'It seems quite impossible to them that a woman is not there.'

- *Dire* is listed here as three separate governors (*dire* (infinitive) 'to say', *pouvoir dire* 'can say', and *vouloir dire* 'want to say') in order to avoid potential skewing of the results with the non-applications of all of the other occurrences of the frequent verb *dire*. In addition, the rate of subjunctive varies in each of these contexts, between 6% (*vouloir dire*) and 42% (*pouvoir dire*). An example of the subjunctive with *dire* and *pouvoir dire* each is given below, in (e-f).
- Of the 49 verbal governors, the construction *être X (adj/noun) que* is counted as a single governor despite the fact that twenty-two different lexical items appear in this context: *bon* 'good', *content* 'happy', *dommage* 'unfortunate', *important* 'important', *incontestable* 'incontestable', *intraitable* 'inflexible', *logique* 'logical', *nécessaire* 'necessary', *normale* 'normal', *obligatoire* 'obligatory', *odieux* 'horrible', *pas la peine* 'not worth it', *mauvais* 'bad', *possible* 'possible', *préférable* 'preferable', *le principe* 'the principle', *question de* 'question of', *rare* 'rare', *ravi* 'overjoyed', *sûr* 'sure', *temps* 'time', *valable* 'valid'. The main reason for this decision is that the vast majority of these constructions behaved in a nearly identical manner (i.e., all but one of them exhibit subjunctive rates of 100% across one to six tokens). The only exception is *être sûr que* (neg.), which only co-occurs with the subjunctive 25% of the time (n= 3/12).

	HF				QF			
	Governor	Rate subj. (%)	Governor	Rate subj. (%)	Governor	Rate subj. (%)	Governor	Rate subj. (%)
10 + occurrences (high)	*falloir*	98	*admettre*	52	*falloir*	89	*être +attribute*	23
	être +attribute	84	*aimer (mieux/bien)*	95	*vouloir*	91	*attendre*	85
	vouloir	96	*ce qui fait*	6	*penser* (neg.)	13	*croire* (neg.)	13
	penser (neg.)	72	*c'est pas*	22	*aimer (mieux/bien)*	67	*admettre*	4
	voir	7	*vouloir dire*	6	*espérer*	21	*considérer*	8
	croire (neg.)	72	*faire*	17	*arriver*	11	*empêcher*	8
	pouvoir dire (neg.)	42	*valoir mieux*	100	*avoir/être peur*	64		
	comprendre	50	*préférer*	100				
	sembler	9						
2-9 occurrences (medium)	*faire attention*	14	*éviter*	100	*s'attendre*	13	*(se) souvenir*	100
	supposer	29	*avoir/être peur*	100	*accepter*	33	*avoir hâte*	100
	attendre	100	*souhaiter*	100	*demander*	100	*permettre*	100
	trouver (neg.)	50	*permettre*	50	*souhaiter*	40	*tenir à*	100
	dire	20	*avoir (x) chance*	50	*(se) rappeler* (neg.)	20	*comprendre* (neg.)	50
	suffire	80	*craindre*	100	*(s')étonner*	60	*voir* (neg.)	100
	demander	100	*faire changement à ce*	100	*préférer*	100	*arranger*	50
	tenir à	100	*faire du mal*	50	*dire*	100	*entendre*	50
	se pouvoir	67	*regretter*	100				
	faire en sorte	100	*surprendre*	100				
	avoir de raisons (neg.)	100						
1 occurrence (nonce)	*arriver*	100	*exiger*	100	*proposer*	100		
	s'étonner	100	*faire plaisir*	100	*concevoir*	100		
	arranger	100	*paraître impensable*	100	*désirer*	100		
	avoir à coeur	100	*rendre*	100	*forcer*	100		
	avoir l'impression (neg.)	100	*trouver logique*	100	*surveiller*	100		
	envisager	100						

Table 3.18: Rate of subjunctive under verbal governors in 20th century HF and QF according to overall frequency of governor (in descending order of frequency; shading indicates shared governors across the two varieties).

In order to understand the importance of each governor in system, the measures of subjunctive productivity elaborated by Poplack (2001) and Poplack, Lealess & Dion (2013) were applied in this and in subsequent analyses.

A first important finding is the prominent position of *falloir* in relation to all other verbal governors in HF. While *falloir* has been identified as counting amongst the most frequent governors in some HF studies (e.g., Bieler 1972; Gougenheim et al., 1964), its proportional importance relative to the other verbal governors in HF has not been highlighted. Based on the comparison of *falloir* to the other verbal governors (Table 3.19), it is clear that the former not only has near-categorical associations with the subjunctive (98%), it also accounts for a large proportion of the data. It represents about one third (30%) of all of the verbal governors that co-occur with an embedded subjunctive in HF. Since it is highly frequent and strongly associated with the subjunctive, it also accounts for nearly half of all of the subjunctive morphology found under verbal governors (42%).

	HF				**QF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
GOVERNOR								
falloir	30	98	216/221	42	65	89	1493/1669	76
others	70	60	304/511	58	35	51	460/900	24
Total	100	71	520/732	100	100	76	1953/2569	100

Table 3.19: Importance of *falloir* vs. other verbal governors as shown by each of the verbal governor's portion of the verbal governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th century HF and QF[48].

Poplack (1990) was one of the first to note the proportional importance of *falloir* to subjunctive selection in her analysis of QF. However, in QF, *falloir* plays a much more important

[48] Results for QF adapted from Poplack, Lealess & Dion (2013:166).

role relative to the other verbal governors than it does in HF (Table 3.19). It is strongly associated with subjunctive use (89%-98%) in both varieties, but in QF it also accounts for 65% of the governor pool and, therefore, 76% of all subjunctive morphology! In other words, there is a high rate of co-occurrence with the subjunctive under *falloir* in both varieties, but it represents a much larger proportion (i.e., double or nearly so) of both the governor and subjunctive morphology pool in QF than in HF. Once *falloir* is removed from the equation (Figure 3.2), the rate of subjunctive is actually higher in HF, which is the opposite of the effect noted when all of the governors were combined.

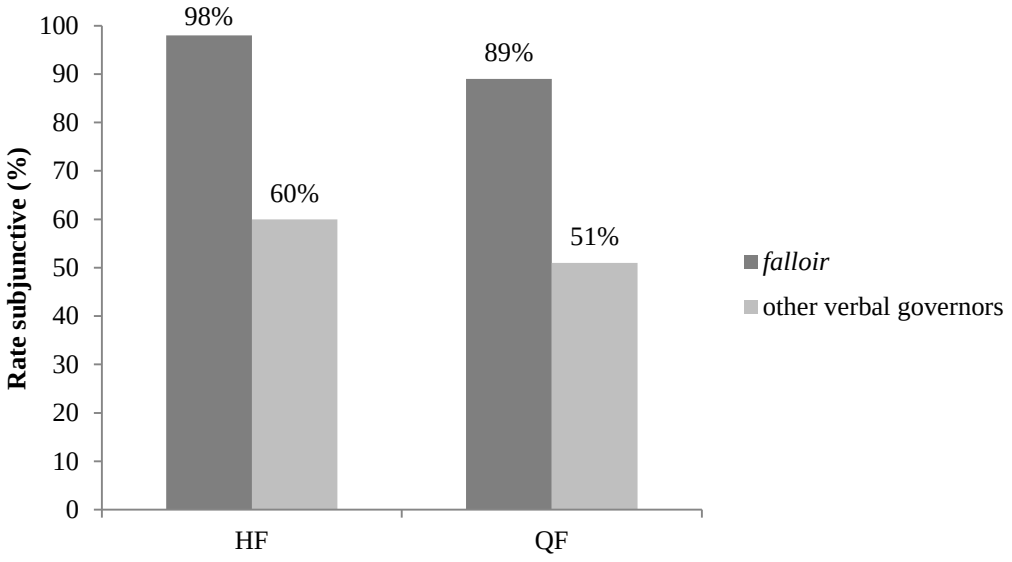


Figure 3.2: Overall rate of subjunctive under *falloir* and other verbal governors in 20th century HF and QF[49].

In addition to *falloir*, Poplack, Lealess & Dion (2013) identified *aimer* and *vouloir* as 'top-tier verbal governors'. In other words, these three verbal governors (all with high rates of co-occurrence with the subjunctive), when combined, represent the majority (80%) of the governor pool and, therefore, nearly all of the subjunctive morphology (92%) found under verbal

[49] Results for QF adapted from Poplack, Lealess & Dion (2013: 166).

governors (Table 3.20). These findings indicate that subjunctive productivity in QF is severely limited under verbal governors given that almost all of the subjunctive morphology is found under only a few. A similar effect is found in HF; however, the top[50] verbal governors that account for the majority of the data also include *penser* (neg.) and *croire* (neg.) in addition to *aimer, vouloir* and *falloir*.

	20th century HF				**20th century QF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
		%	**n**			**%**	**n**	
GOVERNOR								
falloir	30	98	216/221	42	65	89	1493/1669	76
vouloir	8	96	54/56	10	11	91	249/273	13
aimer	3	95	18/19	3	4	67	66/99	3
penser (neg.)	7	72	36/50	7	/	/	/	/
croire (neg.)	5	72	26/36	5	/	/	/	/
Total (top verbal governors)	53	92	350/382	67	80	89	1808/2041	92
Other verbal governors	47	49	170/350	33	20	27	145/528	8
Total	100	69	520/732	100	100	76	1953/2569	100

Table 3.20: Importance of top-tier verbal governors vs. other verbal governors as shown by each of the verbal governor's portion of the governor pool (% pool) and the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th century HF and QF[51].

The five top verbal governors in HF show very high rates of subjunctive. Their individual associations with the subjunctive are actually higher for the most part than in QF and when combined, they account for two thirds of all of the subjunctive morphology (67%), and although this disproportion is striking, it remains quite a bit lower than what was found in QF (92%). Despite the similarities between the two varieties when it comes to proportional disparities across verbal governors, it appears that the extent of this effect is much more pronounced in QF.

[50] Defined here as accounting for at least 3% of the total subjunctive morphology under verbal governors and exhibiting a strong (+70%) association with the subjunctive.
[51] Results for QF adapted from Poplack, Lealess & Dion (2013: 166).

As mentioned above, this difference is largely due to the fact that the preponderance of *falloir* is not echoed to the same extent in HF. In Chapter 4, it will be shown that this is because *falloir* tends to occur more often in HF with an infinitive instead of a subjunctive (or other tense), which may help to explain its relatively low proportion of the data pool and the subjunctive morphology pool. An example of each of these contexts is given below in Table 3.21. If this is an effect that has increased over time, this could be in line with claims that an infinitival complement is being used as a subjunctive replacement strategy (e.g., Bauche 1929). However, it should be noted that, while this would theoretically be possible for most impersonal governors (e.g., one exception is *il se peut* 'it is possible', cited in Bieler 1972), it is not possible for the vast majority unless the main and embedded subject are the same (Harris 1978:226). In other words, an infinitival complement is not possible in all contexts in which the subjunctive can be used and, therefore, does not qualify as a variant.

Context	**Example**
Falloir +INF	« Quand on allait en Turquie, fallait partir (INF) avec ses deux roues de secours. » (ESLO2/MC59/664) 'When we went to Turkey, we had to leave with both of our emergency tires.'
Falloir que +FIN	« Il fallait qu'il ***parte*** (S) euh d'urgence. » (ESLO2/PT45/123) 'He had to leave um urgently.'

Table 3.21: Examples of *falloir que* vs. f*alloir INF* in HF.

Summarizing, five verbal governors (*falloir, aimer, vouloir, croire* (neg.), *penser* (neg.)) exhibit strong associations (i.e., rates) with the subjunctive in HF which, when combined, account for more than half of the entire pool of verbal governors (53%) and, therefore, the majority (67%) of all of the subjunctive morphology. This is strongly reminiscent of the QF

result, though somewhat diluted, where Poplack, Lealess & Dion (2013) found three verbal governors (all of which appear in HF's top five: *falloir, aimer, vouloir*) to account for a much larger proportion (92%) of the subjunctive morphology than is the case in HF. But the main difference between the two varieties resides in the propensity to use a single one of those governors: *falloir*. The overwhelming preponderance of *falloir* in QF helps to explain why the overall rate of subjunctive under verbal governors is lower in HF (71%) compared with QF (76%).

Conditioning of variant choice

Since rates alone are not always a reliable measure of (dis)similarities between two varieties or cohorts, the internal and external factors, detailed in Section 3.3.3 above, were examined in order to determine their relative contribution to subjunctive use under verbal governors in HF using a multivariate analysis. This method will identify the factors that best account for the variability by considering all of the potential conditioning factors simultaneously.

Semantic class

As was mentioned in Section 3.2.2, some previous studies of subjunctive use in HF have reported rates of occurrence according to semantic class (e.g., Lagerqvist 2003; Sand 2003; 1983), although the classification of the different semantic classes can vary from one study to the next. A general finding in the studies that contrasted rates of subjunctive under verbs of volition, emotion, and opinion is that in the first two categories, subjunctive use was found to be largely categorical, while under verbs of opinion, there was variability, which is often interpreted semantically.

Results from the present analysis show that verbal governors that fall into the categories of emotion, necessity, and volition co-occur with the subjunctive at a higher rate than those which fall under the categories of opinion/possibility as well as those which were not easily classified into one of the commonly cited categories ('other') (Table 3.22).

	%	**N**
Emotive	100	19/19
Necessity	98	218/223
Volition	77	95/124
Opinion/possibility	55	136/246
Other	43	52/120

Table 3.22: Rate of subjunctive according to semantic class in 20th century HF.

At first glance, it would appear that these results bear resemblance to what has been reported in HF to date. However, it is argued here that semantic class categories shed little light on subjunctive use in contemporary HF for two reasons. First, there is the presence of underlying lexical effects. For example, the categories of volition and necessity are dominated by *vouloir* and *falloir* respectively, which account for the majority of the data and are independently associated with high rates of subjunctive. *Falloir* accounts for 99% (n= 216/218) of the subjunctive forms found under verbs of necessity, while *vouloir* accounts for more than half (56%, n= 54/95) of those found under verbs of volition. Poplack (1992; 1990) and Poplack, Lealess & Dion (2013) found similar lexical effects in their study of QF, in particular that *vouloir* dominated the volition category, while *aimer* accounted for most of the subjunctive occurrences within the emotive class of verbs.

In addition, it is worth noting that within the opinion category, there is a very strong association (97%, n= 35/36) with what Poplack, Lealess & Dion (2013) identify as 'evaluative verbs', which essentially refers to the construction *être +attribute*. When these contexts are removed from the opinion category, the rate of subjunctive falls slightly to 48% (n=100/209).

With regards to the emotive category, nearly half of the limited number of occurrences (n= 8/19) consist of *aimer* and *avoir peur*, both of which exhibit independent strong associations with the subjunctive. Other quantitative studies of HF rarely provide rates for each individual governor within a semantic class (e.g., Lagerqvist 2003); therefore, they have not reported the link between high subjunctive rates under a particular category and lexical disproportions within them.

Second, some verbal governors exhibit both high and low rates of subjunctive within the same semantic class. For example, *croire* (neg.) and *supposer*[52], both within the category of opinion verbs, exhibit vastly different rates of subjunctive: 72% (n= 26/36) and 29% (n= 2/7) respectively. Such results cast doubt on the ability of semantic class to accurately predict subjunctive use, because if this were a reliable measure, one would expect all members of the class to behave similarly.

Based on clear lexical effects within semantic classes and disparate associations with subjunctive selection, this factor group is not included in the multivariate analysis presented below.

Multivariate analysis

Following Poplack, Lealess & Dion (2013), *falloir* is analyzed separately from the other verbal governors due to its proportional importance. While its disproportionate representation in relation to the other verbal governors is much greater in QF than in HF, the motivation here is to enhance comparability as much as possible between the two studies.

[52] These consist only of tokens of *supposer* 'assume/suppose' and not *supposons* 'let's assume/suppose' as the subjunctive was not found to co-occur with the latter.

Poplack, Lealess & Dion (2013) found an overall rate of 89% subjunctive under *falloir* in QF, and that in the rare cases where it is not used, its absence is best explained by structural (not semantic) factors: conditional matrices and regular (as opposed to suppletive) embedded verbs disfavour the subjunctive. Both effects will be further discussed in the ensuing sections.

In HF, on the other hand, the rate of subjunctive under *falloir* is higher (98%) than it is in QF, but this governor represents a much smaller portion of the governors admitting a subjunctive. In fact, the absence of the subjunctive under *falloir* is only noted five times in the 20th century HF sample, therefore a multivariate analysis is not conducted on these data. It is interesting to note, however, that none of these five occurrences were uttered by the youngest group of speakers (aged 18-24), and the vast majority (n=4/5) were uttered by male speakers, suggesting that subjunctive use in this context has strengthened in apparent time.

Other verbal governors

All other verbal governors (other than *falloir*) were combined[53] and subjected to a multivariate analysis using the same factors analyzed in QF, and then compared to Poplack, Lealess & Dion's (2013) findings.

The results of the analysis are given in Table 3.23 and can be read as follows here, and in subsequent multivariate analyses. The *corrected mean* is considered the "average frequency of occurrence of the application value of the dependent variable" (Paolillo 2002:79), the application value of the dependent variable in this case being the use of the subjunctive over any other embedded tense. *Factor weights* (abbreviated to FW in the tables) reflect the probability that a

[53] While four other verbal governors (in addition to *falloir*) account for more of the subjunctive morphology than the others, they are included here under 'other verbal governors'. This method of analysis is also in line with that of (Poplack, Lealess & Dion 2013) and, therefore, it maintains comparability between the two studies.

particular factor has an effect on the use of a specific variant (i.e., the subjunctive). They can have a value between 0 and 1, with values closer to 0 considered to disfavour the choice of the variant, while values closer to 1 favour the selection of that variant. Some analysts use the 0.5 mark to distinguish between favouring and disfavouring factors, while others suggest that it is more important to consider the factor weights of each factor in relation to each other (Tagliamonte 2006:141,145,156). Only factor groups that are selected as having a statistically significant effect on variant choice are presented in this and subsequent tables. When more than one factor group is selected as significant in a multivariate analysis, it is possible to determine which has the strongest effect by use of the *range*. The range (also referred to as the *magnitude of effect*) can be described as the difference between the highest and lowest factor weights in any factor group selected as statistically significant (Tagliamonte 2006:242). The higher the range (expressed as a whole number), the stronger the effect. Therefore, if two factor groups are selected as statistically significant and the range is higher in one, the analyst can conclude that the factor group with the higher range has the strongest effect on variant choice relative to the other factor group (Tagliamonte 2006:242).

Based on the results presented in Table 3.23, a first observation is that in HF, two of the three factor groups selected as statistically significant are structural: tense of matrix verb and presence of *que*. The presence of complementizer *que* plays an important role in subjunctive selection under verbal governors other than *falloir* in HF. This factor group was selected as statistically significant in QF as well and exhibits the same direction of effect (i.e., its presence highly favours subjunctive use).

	HF 20th C			QF 20th C[54]		
Corrected mean	*.594*			*.526*		
Total N	511			900		
	FW	%	N	FW	%	N
SEMANTIC						
Sentence Type[55]						
Affirmative	.46	55	181/327	[ ]		
Negative	.57	64	105/163	[ ]		
	11					
Other indicators of non-factual modality[56]						
Yes	[ ]	54	45/83	[ ]		
None	[ ]	58	227/391	[ ]		
MORPHOSYNTACTIC						
Tense						
Conditional	.86	91	30/33	.25	46	55/119
Imperfect	.67	76	22/29	.65	66	94/142
Present Indicative	.46	56	236/419	.51	49	271/555
Passé composé	.44	56	9/16	.42	39	14/36
Future	.38	50	3/6			
Range	*48*			*40*		
Distance between governor and embedded verb						
None	[ ]	62	175/284	[ ]		
Syntactic	[ ]	57	74/130	[ ]		
Parenthetical	[ ]	57	55/97	[ ]		
Range						
Presence of *que*						
Yes	.54	63	301/481	.52	52	395/756
No	.09	10	3/30	.39	45	64/141
Range	*45*			*13*		
Morphological form of embedded verb						
Suppletive	[ ]	58	245/423	.56	53	340/639
Regular	[ ]	67	59/88	.36	47	120/253
Range				*20*		

KO indicates 'knock out' (i.e., there is no variability with a specific factor), here because the subjunctive occurs categorically (100%); [ ] denotes factors which were not selected as statistically significant; n/a denotes factor groups which were not tested; Highlighting indicates similarities across varieties

Table 3.23: Multivariate analysis of factors contributing to subjunctive use under verbal governors other than *falloir* in 20th century HF and QF (shading indicates shared effects across varieties.)

[54] The marginal results are not available for QF in cases where the factor group was not selected as statistically significant.

[55] Only affirmative and negative are considered under sentence type in order to avoid interactions and skewing of the results. Conditional *si*-clauses are already included under 'other indicators of non-factual modality), and 95% of the interrogatives sentences are governed by *vouloir* (n= 18/19).

[56] In order to ensure that the factor groups remain independent of one another, the tokens that were coded as having tense as the only indicator of non-factual modality were excluded due to an interaction with tense of the main clause.

The tense of the matrix verb exhibits a significant effect on subjunctive use in both varieties; however, there are some differences to note here. Starting with HF, both conditional and imperfect matrix verbs favour the use of the subjunctive in the embedded clause (Table 3.23).

Both future and *passé composé* matrix clauses strongly disfavour the subjunctive, while the present indicative does so mildly. Interestingly, the vast majority of indicative tokens occurring under a future or *passé composé* matrix were governed by *voir* and *comprendre* (80%, n= 8/10), which further points to the importance of the lexical identity of the governor. While the propensity of the subjunctive to co-occur with conditional matrices may, at first glance, lend support to the semantic hypothesis, further analysis of the tokens that make up the subjunctive-favouring environments (i.e., both conditional and imperfect matrix verbs) in HF revealed that nearly half (n= 14/30) of conditional verbal governors which co-occurred with a subjunctive were *vouloir* and *aimer,* both of which are independently associated with high rates of subjunctive. Similarly, the same two verbal governors represent two thirds (n= 14/22) of the imperfect matrices that co-occurred with a subjunctive. These findings nullify the apparent effect and suggest instead that subjunctive use according to matrix tense is lexically motivated.

In QF, on the other hand, Poplack, Lealess & Dion (2013) found the same effect with the other verbal governors as they did with *falloir,* namely that the subjunctive is strongly disfavoured with conditional matrices (in contrast to HF), while it is favoured with imperfect and present indicative ones. Poplack (1992) showed that conditionals (rather than subjunctives) were more likely to be embedded under conditional matrices and attributed this finding to a "conditional concordance effect" (ibid., 248). Since no such disfavouring effect of subjunctives embedded under conditional matrices was found here, it can be concluded that this tense effect is

not operative in HF. In fact, there is only a single token in the entire data set where a conditional matrix selects a conditional embedded verb (15).

(15) « Je voudrais dire qu'il faudrait leur apprendre un petit peu de ce qu'on ne fait pas du tout en France d'instruction civique. » (ESLO1/1268/645)
'I would like to say that we should teach them a little bit of what we don't do at all in France: civic education.'

Finally, neither the morphological form of the embedded verb nor the presence of intervening material contributes significantly to subjunctive selection in HF under verbal governors other than *falloir*, although the rates alone suggest that contexts in which such material is absent are more favourable to subjunctive use over those in which some is present. With regards to the morphological form of the embedded verb, although not selected as a significant contributing factor to subjunctive use in HF, the direction of effect shows that the subjunctive is more likely to co-occur with regular embedded verbs as opposed to suppletive ones, which is the opposite of what was found (and selected as statistically significant) in QF. This will be discussed further in the following sections.

Moving on to the semantic factor groups, only one has a significant effect on subjunctive use in HF: sentence type. It appears that the subjunctive is more likely to be used when the matrix clause is negated, which partially supports the semantic hypothesis. However, closer examination of this context revealed that the majority of the subjunctive tokens (59%, n= 62/105) in the negative category consist of two of the top governors: *croire* (neg.) and *penser* (neg.), both of which only co-occur with the subjunctive when the governor is negated. In fact, once these two governors are removed from the total, the rate of subjunctive under negated

verbal governors is 57% (n= 43/75), which is almost identical to the rate found under affirmative ones (55%, n= 181/327). In other words, this apparent semantic effect is masking a lexical one.

The other semantic factor group, presence of other indicators of non-factual modality, has no effect at all. Recall that this factor group was included in order to test the idea that the presence of certain elements in the sentence (e.g., a conditional *si*-clause) could help to convey some of the semantic readings typically associated with the subjunctive. If the subjunctive is, in fact, semantically motivated, one could expect it to be found without restriction in contexts in which such indicators are also found. These results suggest that this is not the case. Not only is this factor group not selected as significant, the direction of effect of the marginal results shows that the subjunctive is actually more likely to be used in contexts devoid of such indicators.

In QF, Poplack, Lealess & Dion (2013) found semantic class to be the only semantic factor group selected as statistically significant; however, they were able to successfully show that this result was the result of underlying lexical factors, similar to what was shown above for HF.

Summary

In this section, the overall rate and underlying conditioning of subjunctive use in HF was examined in detail. These findings were then compared to those of Poplack, Lealess & Dion (2013) for QF. The overall rate of subjunctive under verbal governors in 20th C HF (71%) is nowhere near prescriptive ideals and is also *lower* than the overall rate in QF (76%), contrary to expectations. Despite what appears at first glance to be a statistically significant difference in terms of overall rates between the two varieties, this difference is largely attributable to the lower proportion of *falloir* in HF compared to QF.

Subjunctive morphology was found under a total of 49 verbal governors in HF compared with only 34 in QF. However, further analysis revealed that only a small portion of these co-occur frequently in the data (10+ times) and that subjunctive rates vary drastically regardless of governor frequency. Five verbal governors (*aimer, vouloir, croire* (neg.), *penser* (neg.), and *falloir*) play a greater role than the others in that, when combined, they represent half of the governor pool and two thirds of all of the subjunctive morphology. While the dispersion of the subjunctive over a greater number of contexts in HF may suggest greater productivity, this is countered by the fact that the rate of subjunctive is much lower (49%) under verbs other than the top five, where it is nearly categorical (92%), as shown in above in Table 3.20.

In terms of internal conditioning, the results of the multivariate analyses show that semantic factor groups do not explain the variability of the subjunctive under verbal governors in HF. In the one case where a semantic factor group appeared to have explanatory power at first glance, it was shown that it was, in fact, due to an underlying lexical effect. This finding closely mirrors that of Poplack, Lealess & Dion (2013) for QF.

The only factors that were found to play a significant role in the conditioning of subjunctive use in HF were morphosyntactic (presence of *que* and tense of matrix verb). In other words, subjunctive morphology is most likely in contexts where complementizer *que* is present, and with conditional and imperfect matrices, the latter result being attributable to lexical effects. Note that absence of intervening material, although not selected as significant, appears to favour subjunctive use as well.

In QF, Poplack, Lealess & Dion (2013) found that suppletive embedded verbs favoured subjunctive use under all verbal governors. The presence of intervening material between the

governor and the embedded verb favoured subjunctive selection under *falloir*, while under the other verbal governors it was more so the presence of the complementizer *que*. The selection of matrix tense as significant under all verbal governors was attributed to the remnants of a tense concordance effect, which resulted in a strong disfavouring effect of the subjunctive under conditional matrices.

It would appear that in both varieties of French, subjunctive use has limited productivity as evidenced by a) a lack of semantic motivation, b) rate discrepancies across contexts that govern the subjunctive and c) the fact that morphosyntactic and lexical conditioning factors best account for variant choice. While such effects appear to be stronger in QF, they are also clearly present in HF as well.

3.4.1.2 Non-Verbal governors

Overall rates

The overall rate of subjunctive under non-verbal governors in HF is 73%, which is roughly the same rate as what was found under verbal governors (70%) (Table 3.24).

HF	QF
73% (n= 106/145)	82% (n= 163/198)

Table 3.24: Overall rate of subjunctive under non-verbal governors in 20th century HF and QF[57].

When comparing the rates of subjunctive under non-verbal governors in HF and QF, it is clear that in HF, the rate is once again *lower* than what was found by Poplack, Lealess & Dion, (2013) for QF (82%). This difference is statistically significant according to Fisher's exact test (p= 0.0465). Just like with the verbal governors, this finding is contrary to the general

[57] Results for QF are adapted from Poplack, Lealess & Dion (2013).

expectation that HF will exhibit more subjunctive than QF. Further analysis will reveal whether the overall rate difference between the two varieties is actually due to structural differences or whether it is more superficial in nature.

Lexical identity of non-verbal governors

Subjunctive morphology appears under a total of 23 non-verbal governors in 20th century HF, which is substantially lower than the number of verbal governors (n= 49). But just like with the verbal ones, some of the non-verbal governors exhibit high rates of co-occurrence with the subjunctive, while others exhibit low ones, as is shown in Figure 3.3. In addition, there are clear differences in the frequency of the individual governors: only a small number of them (17%, n= 4/23) are frequent (i.e., 10+ occurrences). Among the most frequent ones, the rate of subjunctive varies from 13% to 100%, but most (n= 3/4) exhibit high rates of subjunctive (71%-100%). Similar to the pattern with verbal governors, most of the high subjunctive governors are low-medium frequency. In other words, the subjunctive does not freely occur in all eligible contexts (frequent or otherwise), which signals a lack of productivity.

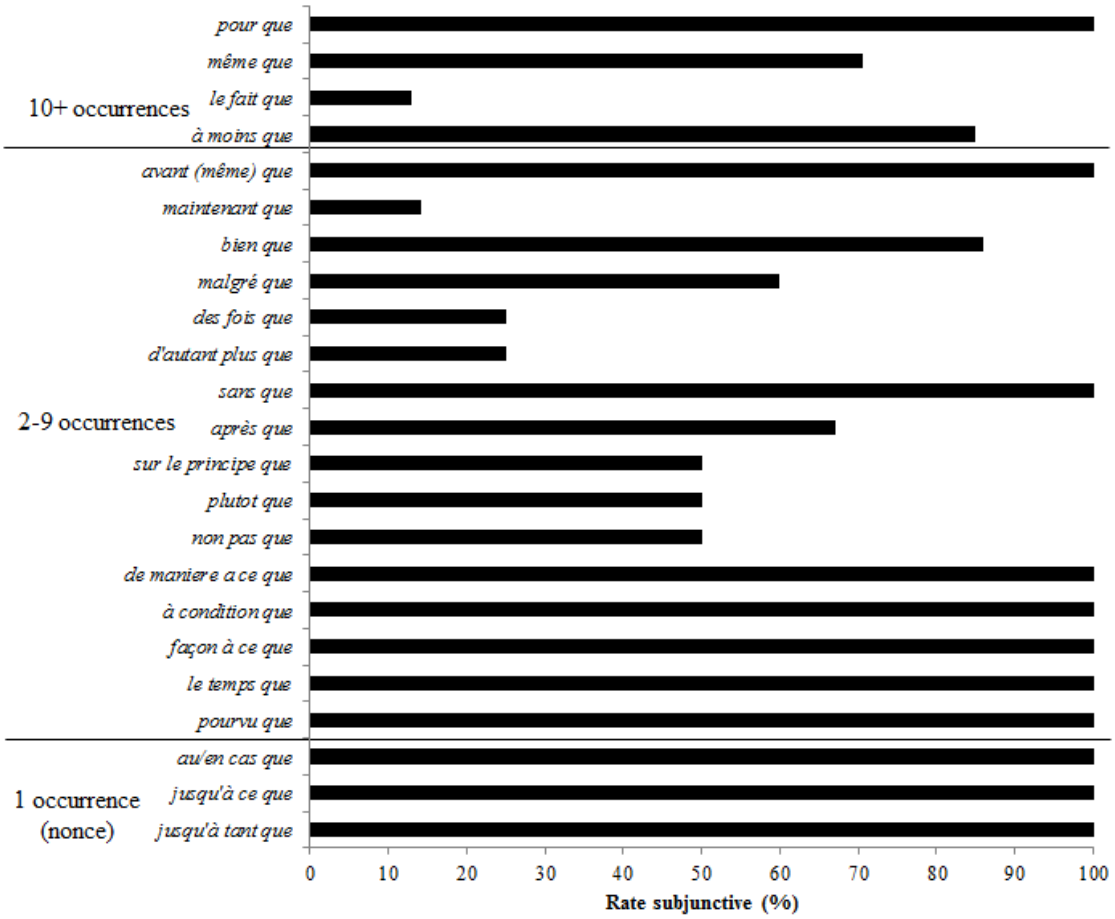


Figure 3.3: Rate of subjunctive under each individual verbal governor in 20th century HF according to overall number of occurrences of the governor (listed in descending order of frequency).

The subjunctive co-occurs with slightly more non-verbal governors in HF (n=23) than in QF (n= 14) (Poplack, Lealess & Dion, 2013), but several of these (n= 9) are shared across the two varieties (Table 3.25). In fact, nearly two thirds of the non-verbal governors identified in QF (64%, n=9/14) also occur in HF, while the rest can be attributed to cross-variety lexical differences. In addition, the vast majority of frequent governors exhibit high rates of subjunctive, which further confirms Bybee & Thompson's (1997) Conserving Effect. Despite some lexical differences, the same general pattern emerges in both varieties (as was the case for verbal governors): only a small number of governors exhibit both high rates of subjunctive and high

overall frequencies. The relative contribution of these to overall subjunctive use is explored in more detail below.

	HF		QF	
	Governor	Rate subj. (%)	Governor	Rate subj. (%)
10 + occurrences (high)	*pour que*	100	*pour que*	98
	même que	71	*mais que*	91
	le fait que	13	*avant (même) que*	87
	à moins que	85	*soit que*	50
			jusqu'à tant que	82
			à moins que	70
2-9 occurrences (medium)	*avant (même) que*	100	*sans que*	100
	bien que	86	*le temps que*	60
	maintenant que	14	*à (en) ce que*	33
	malgré que	60	*au/en cas que*	67
	sans que	100	*jusqu'à ce que*	100
	d'autant plus que	25	*en tout cas que*	50
	des fois que	25	*jusqu'à que*	50
	après que	67		
	pourvu que	100		
	le temps que	100		
	façon à ce que	100		
	à condition que	100		
	de manière à ce que	100		
	non pas que	50		
1 occurrence (nonce)	*jusqu'à tant que*	100	*façon à ce que*	100
	jusqu'à ce que	100		
	au/en cas que	100		

Table 3.25: Rate of subjunctive under non-verbal governors in 20th century HF and QF according to overall frequency of governor (shading indicates shared governors across the two varieties)[58].

Of the non-verbal governors, one in particular is far more frequent than the others: *pour que* (Table 3.26). The use of Poplack's three measures of productivity revealed that *pour que* alone accounts for one quarter of the data pool (26%) and more than one third of the subjunctive morphology (36%). Its distribution mirrors that of *falloir que* under the verbal governors very

[58] The translation for each of the verbal governors found in HF and QF is given in Appendix A.

closely. In addition, the distribution of *pour que* in HF is nearly identical to that in QF. When *pour que* is separated from the other non-verbal governors, the overall rate of subjunctive drops to 64% in HF. While this study is not the first to show the strong association of *pour que* with the subjunctive in HF (e.g., Sand 2003) or its frequency (e.g., Gougenheim et al. 1964), no other study to date has identified its proportional importance with regards to subjunctive use under non-verbal governors.

	HF				**QF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
pour que	26	100	38/38	36	27	98	53/54	33
Other governors	74	64	68/107	64	73	76	110/144	67
Overall total	100	73	106/145	100	100	82	163/198	100

Table 3.26: Importance of *pour que* vs. other non-verbal governors as shown by each of the non-verbal governor's portion of the governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th century HF and QF.

Pour que is not the only non-verbal governor to occupy a more prominent position relative to the others in HF. Table 3.27 displays the other top non-verbal governors that also have very high associations with the subjunctive, while occupying a notable proportion of all of the subjunctive morphology[59]. When these governors are combined, the rate of subjunctive with them is nearly categorical (96%) and nearly double that of the other non-verbal governors (51%). In addition, they account for nearly two thirds (64%) of all of the subjunctive morphology found under non-verbal governors. In other words, subjunctive use under non-verbal governors in HF is largely lexically conditioned in that it is mostly found under a handful of top governors. Under the other non-verbal governors, the indicative is just as likely to be used as the subjunctive.

[59] Defined here as having a rate of co-occurrence with the subjunctive of 80% or more and occupying 4% or more of the subjunctive morphology pool.

	% Pool	Rate Subj. %	n	% Subj.
GOVERNOR				
pour que	26	100	38/38	36
à moins que	9	85	11/13	10
bien que	5	86	6/7	6
avant que	6	100	9/9	8
sans que	3	100	4/4	4
Total top-tier	49	96	68/71	64
Other governors	51	51	38/74	36
Overall total	100	73	106/145	100

Table 3.27: Importance of top-tier non-verbal governors vs. other non-verbal governors as shown by each of the non-verbal governor's portion of the non-verbal governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th century HF.

The disproportionate pattern of subjunctive use under non-verbal governors in HF corresponds closely to that reported for QF by Poplack, Lealess & Dion (2013) and St-Amand (2002). They found that a handful of governors accounted for nearly all of the subjunctive morphology (88%) in QF, which is substantially higher than in HF (64%) (Table 3.28).

	% Pool	Rate Subj. %	n	% Subj.
GOVERNOR				
pour que	27	98	53/54	33
avant que	16	87	27/31	17
mais que	18	91	32/35	20
jusqu'à tant que	9	82	14/17	9
soit que	14	50	14/28	9
Total top-tier	84	85	140/165	88
Other governors	16	70	23/33	12
Overall total	100	82	163/198	100

Table 3.28: Importance of top-tier non-verbal governors vs. other non-verbal governors as shown by each of the non-verbal governor's portion of the non-verbal governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th century QF[60].

[60] Adapted from Poplack, Lealess & Dion (2013:176).

Comparing the HF and QF results, it is apparent that the same kind of lexical disproportions exist in both varieties, even if the exact governors involved are not always the same[61]. *Pour que* (as discussed above) patterns nearly identically in both varieties, accounting for the largest proportion of the non-verbal governors and the largest share of subjunctive morphology, while *avant que* exhibits stronger associations with the subjunctive in HF (100%) than QF (87%) but accounts for a smaller portion of both the data pool and the subjunctive morphology pool (i.e., it occurs less in the HF data than it does in the QF data). The other top governors identified in Tables 3.27 and 3.28 differ across the two varieties, but in both cases, they account for the bulk of the governors and the bulk of the subjunctive morphology[62]. Just like with the verbal governors, the top tier in HF occupies a less important role proportionally than in QF, thus resulting in the subjunctive being spread across more governors in HF, hinting at a higher (although still quite limited) degree of productivity.

Conditioning of variant choice

Several internal (linguistic) factors were analyzed in order to determine their relative contribution to subjunctive use with non-verbal governors in HF. The results are then compared with Poplack's findings for QF. In order to test the commonly cited semantic conditioning of subjunctive use, two semantically-defined factor groups are included in this analysis: reality of the predication (irrealis vs. realis) and sentence type (affirmative vs. negative). If there is a semantic effect conditioning subjunctive use, higher, if not categorical, rates of subjunctive in

[61] Such a finding is not unusual. Poplack et al. (in press) found in their study of the subjunctive across several Romance languages, that lexical disproportions, although present in each, were language-specific.
[62] In fact, two of the top non-verbal governors identified by Poplack, Lealess & Dion (2013) do occur sparingly in the HF corpus but do not co-occur with the subjunctive at all (*mais que* and *soit que*).

irrealis and negative contexts would be expected. At the very least, the semantic factors should outweigh the morphosyntactic ones in accounting for subjunctive selection.

	HF 20th C			QF 20th C[63]		
Corrected mean	*.794*			*.833*		
Total N	145			198		
	FW	%	N	FW	%	N
SEMANTIC						
Reality of the predication						
Irrealis	.72	90	69/77	[ ]		
Realis	.25	52	34/65	[ ]		
Range	*47*					
Sentence Type						
Affirmative	[ ]	73	103/141	[ ]		
Negative	[ ]	67	2/3	[ ]		
Range						
MORPHOSYNTACTIC						
Distance between governor and embedded verb						
None	[ ]	70	57/82	[ ]		
Syntactic	[ ]	80	28/35	[ ]		
Parenthetical	[ ]	75	21/28	[ ]		
Range						
Presence of *que*						
Yes	.52	75	105/141	.53	85	158/186
No	.05	25	1/4	.13	42	5/12
Range	*47*			*40*		
Morphological form of embedded verb						
Suppletive	.43	69	82/119	[ ]		
Regular	.79	92	24/26	[ ]		
Range	*36*					

KO indicates 'knock out' (i.e., there is no variability with a specific factor), here because the subjunctive occurs categorically (100%); [ ] denotes factors which were not selected as statistically significant; n/a denotes factor groups which were not tested; Highlighting indicates similarities across varieties

Table 3.29: Multivariate analysis of factors contributing to subjunctive use under non-verbal governors in 20th century HF and QF[64,65](shading indicates shared effects across varieties.)

Results for the multivariate analysis are presented in Table 3.29. A first observation is that in HF, two of the three factor groups selected as significant in their contribution to

[63] The marginal results are not available for QF in cases where the factor group was not selected as statistically significant.

[64] Despite *pour que*'s proportional disparity compared to the other non-verbal governors, it is included in the multivariate analysis in order to maintain comparability with QF.

[65] Results for QF adapted from Poplack, Lealess & Dion (2013: 182).

subjunctive use are structural: the morphological form of the embedded verb and presence or absence of complementizer *que*. These are two of the same factor groups that play a role in the selection of the subjunctive under verbal governors (though morphological form of the embedded verb was not selected as significant under verbal governors, the direction of effect is the same). The subjunctive is favoured by non-suppletive (i.e., regular) embedded verbs and in contexts in which the complementizer *que* is present (as opposed to absent), the latter effect being stronger (as evidenced by higher range).

One semantic factor group is also selected as statistically significant in HF: the reality of the predication, with irrealis contexts favouring subjunctive use. This would appear at first glance to coincide with the widely accepted idea that the subjunctive expresses a wide range of non-actualized situations. However, the majority (70%, n = 48/69) of the occurrences of the subjunctive in irrealis contexts are constituted of the five non-verbal governors introduced in the previous section, each independently exhibiting strong associations with the subjunctive and occupying a large portion of the data. In QF, Poplack, Lealess & Dion (2013) found only that only one factor had an effect on variant choice: the presence of complementizer *que,* an effect which is shared by HF speakers.

Suppletive vs. regular embedded verbs

A clear difference noted above between the two varieties under non-verbal governors is that HF speakers have higher associations with the subjunctive with regular embedded verbs than with suppletive ones, while in QF, the opposite is true. While this effect was not selected as statistically significant with verbal governors, the direction of effect was the same. In order to investigate this further, a list of all of the suppletive and regular verbs embedded under a subjunctive governor (both verbal and non-verbal) in 20th century HF is given in Table 3.30.

SUPPLETIVE			REGULAR		
Embedded verb	**Rate Subj.**		**Embedded verb**	**Rate Subj.**	
	%	N		%	N
Aller 'to go'	84	36/43	*Dire* 'to say'	77	27/35
Avoir 'to have'	65	122/189	*Apprendre* 'to learn'	73	8/11
Être 'to be'	67	213/318	*Prendre* 'to take'	91	10/11
Faire 'to make'	82	70/85	*Venir* 'to come'	100	10/10
Pouvoir 'to be able'	65	24/37	*Comprendre* 'to understand'	100	10/10
Savoir 'to know'	79	11/14	*Écrire* 'to write'	90	9/10
Vouloir 'to want'	50	4/8	*Sortir* 'to leave'	100	8/8
Falloir 'to be necessary'	17	1/6	*Mettre* 'to put'	83	5/6
			Reprendre 'to resume'	100	4/4
			Plaire 'to please'	80	4/5
			Devoir 'must'	0	0/4
Other (3 tokens or less)	100	2/2	Other (3 tokens or less)	79	48/61
Total	**69**	**483/702**	**Total**	**82**	**143/175**

Table 3.30: List of suppletive and regular verbs in 20th century HF under verbal and non-verbal governors and their respective rates of subjunctive.

Since the rate of subjunctive with suppletive embedded verbs is lower (69%) than what it is with regular embedded verbs (82%) in HF, it means that in HF, speakers are more likely to use an inflectional verbal ending to denote the subjunctive rather than a stem-changing verb. Non-suppletive embedded verbs have a much higher type frequency (i.e., the number of "different lexical items a certain pattern or construction is applicable to" (Bybee & Thompson 1997:378)) than suppletive ones, and can, therefore, be an indicator of high productivity. Stem-changing rules apply to much fewer lexical items and thus these items (i.e., suppletive verbs) have a low type frequency and are more likely to undergo entrenchment, particularly if they exhibit high token frequency.

Table 3.30 clearly shows that the rate of subjunctive with suppletive verbs in HF is inconsistent: some have very high rates, while others exhibit much lower rates. In fact, the two most frequent embedded verbs (*avoir* and *être*) both exhibit fairly low rates of subjunctive and

thus appear to be driving down the overall rate of subjunctive with suppletive verbs. Such inconsistencies within a single factor suggest that the suppletive/regular distinction does not adequately account for the variability, and that lexical effects are likely at play (i.e., subjunctive morphology is simply more likely to be found on specific embedded verbs than on others, despite their morphological status).

Summary

In this section, the overall rate and underlying conditioning of subjunctive use under non-verbal governors in HF has been examined in detail. Results revealed an overall subjunctive rate of 73% under a total of 23 non-verbal governors which is significantly lower than what was found by Poplack, Lealess & Dion (2013) for QF: 82% subjunctive under 14 non-verbal governors. However, just like with verbal governors, the majority of the subjunctive morphology is concentrated under a single non-verbal governor, *pour que,* in both varieties. This finding closely mirrors that of *falloir que* under verbal governors. Four other top non-verbal governors (*avant que, sans que, à moins que, bien que*) were also identified in HF to have strong associations with the subjunctive and, when combined, account for a substantial portion of both the governor and subjunctive morphology pools. With regards to internal conditioning, the multivariate analysis revealed the relative importance of structural factors as opposed to semantic ones, which parallels the situation in QF.

Even though the specific lexical identity of individual governors varies somewhat across the two varieties, the overall pattern of subjunctive use is nearly identical in both, in that a) it is conditioned by structural and lexical factors as opposed to semantic ones and b) a handful of lexical items account for the majority of the data, with *pour que* leading the pack in both cases. This effect has not been previously reported in studies of subjunctive use in HF. Even though the

top-tier non-verbal governors in QF account for a larger proportion of both the governor and subjunctive morphology pools compared to HF, the rate of subjunctive associated with the top non-verbal governors is higher in HF (96%) than in QF (85%), while the rate of subjunctive with non-top-tier non-verbal governors is much lower in HF (51%) than in QF (70%). This helps to explain the overall rate discrepancy across the two varieties and is indicative of a lack of productivity in both, although manifested differently (i.e., strong subjunctive associations across a larger portion of the pool in QF vs. nearly categorical subjunctive associations across a smaller portion of the pool in HF). In addition, the subjunctive is found under fewer non-verbal governors in QF, similar to what was found with the verbal governors.

3.4.1.3 Embedded verbs

The results for HF presented thus far have shown that lexical and structural factors are at play in a speaker's choice to use the subjunctive, particularly that certain verbal and non-verbal governors are more likely to co-occur with the subjunctive than the others, echoing what was found in QF. Poplack, Lealess & Dion (2013) also identified a similar effect of the embedded verb (i.e., the verb that carries the subjunctive morphology), whereby only a handful of highly frequent embedded verbs not only have high associations with the subjunctive but also account for the majority of the subjunctive morphology. Table 3.31 shows that when the top-tier embedded verbs (*être, avoir, aller, faire*[66] are combined (from both verbal and non-verbal governors), they account for the majority of the governor pool and, therefore, the majority of the subjunctive morphology pool in QF, but their rate of subjunctive is no higher with these four verbs than with the other embedded verbs. Poplack, Lealess & Dion (2013) show that when these

[66] Poplack, Lealess & Dion (2013) also include *pouvoir* as a top embedded verb under non-verbal governors. Here, figures are provided for *être*, *avoir*, *aller*, and *faire* only.

select few embedded verbs are used in conjunction with a top-tier governor, the rate of subjunctive is nearly categorical (92%). In other words, subjunctive use in QF is almost guaranteed in constructions that are made up of a) a top-tier governor and b) one of a small class of embedded verbs.

	% Pool	Rate Subj. %	Rate Subj. N	% Subj.
être	27	68	467/684	23
avoir	16	69	284/412	15
faire	14	87	318/365	16
aller	16	86	354/410	18
Total (top-tier embedded verbs)	73	76	1423/1871	72
Other embedded verbs	27	76	530/698	28
Total (all)	100	76	1953/2569	100
Falloir/aimer/vouloir/pour que +être/avoir/faire/aller	92% subjunctive			

Table 3.31: Importance of the lexical identity of embedded verbs under verbal and non-verbal governors as shown by each of the embedded verb's portion of the pool (% pool), the portion of subjunctive morphology found with each (% subj.) and the overall rate of subjunctive with each (Rate subj.) in 20th century QF[67].

Is the same effect present in HF? Under both verbal and non-verbal governors, a total of 57 different embedded verbs were identified in 20th century HF. The same top embedded verbs identified by Poplack, Lealess & Dion (2013) play a greater role than the others: *être, avoir, aller, faire* in that they are the most frequent, representing 73% of the embedded verb pool. The disproportionate distribution of embedded verbs was also noted for HF by Sand (2003) but no further analysis or explanation was given. When combined, they account for nearly three quarters of the subjunctive morphology and also have a fairly strong association with the subjunctive (69%) (Table 3.32), though less so than for the other embedded verbs (77%). This is because amongst the top embedded verbs, *avoir* and *être* are by far the most frequent, yet the rate

[67] Results for QF adapted from Poplack, Lealess & Dion (2013: 179) with additional analysis.

of subjunctive with them (65-67%) is quite a bit lower than with the others (80-84%), which lowers the average rate. When combined with a top governor, the rate of subjunctive is nearly categorical: 92%, identical to QF.

Interestingly, *avoir* and *être* behave similarly in QF as they do in HF (Table 3.31 and 3.32) in that they exhibit much lower rates of subjunctive than the other top embedded verbs, although their association with the subjunctive is slightly stronger than it is in HF. In addition, these two highly frequent embedded verbs account for 58% of the embedded verb pool in HF, while in QF, they only account for 43%, thus contributing to the difference in the overall rate of subjunctive under the top embedded verbs.

		Rate Subj.		%
	% Pool	**%**	**N**	**Subj.**
être	36	67	213/318	34
avoir	22	65	122/189	19
faire	10	80	70/87	11
aller	5	84	36/43	6
Total (top-tier embedded verbs)	73	69	441/637	70
Total other embedded verbs	27	77	185/240	30
Total	100	71	626/877	100
Falloir/aimer/vouloir/penser/croire/pour que +être/avoir/faire/aller	92% subjunctive			

Table 3.32[68]: Importance of the lexical identity of embedded verbs under verbal and non-verbal governors as shown by each of the embedded verb's portion of the pool (% pool), the portion of subjunctive morphology found with each (% subj.) and the overall rate of subjunctive with each (Rate subj.) in 20th century HF.

Moreover, this divergence also helps to explain why suppletive embedded verbs are more likely to co-occur with the subjunctive in QF than they are in HF since the two most frequent embedded verbs (*avoir* and *être*), despite having slightly stronger associations with the subjunctive in QF, account for less of the data thus resulting in a more influential role for the

[68] Both the verbal and non-verbal governors are combined for the analysis of the embedded verb since most important players are common to both data sets.

other two top-tier embedded verbs *(faire* and *aller*), which not only exhibit higher rates of subjunctive, but also represent a much greater portion of the data pool in QF than in HF (30% vs. 15%).

Summary

In this section, the role of the embedded verb in subordinate clauses under both verbal and non-verbal subjunctive governors was analyzed. Just like in QF, the majority of the subjunctive morphology was found on only a few embedded verbs in HF. However, HF speakers exhibit higher rates of subjunctive on embedded verbs *other* than those identified by Poplack for QF. This apparent difference is simply due to a higher proportion of the two most frequent embedded verbs (*avoir* and *être*) in QF, which tend to exhibit lower associations with the subjunctive than the other top embedded verbs *(aller* and *faire*) and many of the other less frequent embedded verbs.

3.4.1.4 Non-governed subjunctives

Subjunctive use in HF is not only limited to subordinate clauses headed by a verbal or non-verbal governor − it can also occur in non-governed contexts. The definition of the variable context adopted in this study includes the presence of a (verbal or non-verbal) governor, therefore non-governed subjunctives were excluded at the outset. However, for the sake of completeness, they are analyzed separately in this section in order to determine to what degree they are hosting the subjunctive in 20th century HF.

Relative clauses

Some studies of HF have noted the use of the subjunctive in relative clauses (e.g., Lagerqvist 2003; Sand 2003), examples of which are given in (16) and (17).

(16) « C'est la seule -seule fête qui ***puisse*** (S) exister à Orléans. » (ESLO1/PW85/648)

'It is the only-only holiday that can exist in Orléans.'

(17) « J'aurais voulu simplement qu'il existe quelques cafés où on ***puisse*** (S) euh fumer. » (CFPP/AM1/904)

'I would have simply liked there to be more cafes where we could um smoke.'

Subjunctive use is actually prescribed in relatives (e.g., Grevisse & Goosse 1989), as long as it is expressing a goal, intention, consequence, doubt, or following a negated, conditional, or interrogative main clause (ibid., 372). Lagerqvist (2003) reported that 14% of all the subjunctive occurrences he analyzed were found in restrictive (e.g., *la fille qui **ait** (S) les cheveux courts* 'the girl who has short hair') or concessive (e.g., *quoi que ce **soit** (S)* 'anything') relative clauses. Blanche-Benveniste et al. (1990) (referencing Lim's results) reported that relative clauses account for only 4% of all subjunctive occurrences, although it is not specified if the type of relative clause was considered. Neither study provides the number of non-occurrences (i.e., relative clauses with non-subjunctive embedded verbs); therefore, the rate of subjunctive in this context remains unknown. Sand (2003) on the other hand, identified a rate of subjunctive of 78% (n= 14/18), but did not specify if the type of relative clause was limited by restrictiveness or relativizer. It would be safe to assume that the extraction was restricted in some way since it is highly unlikely that he would have only found 18 relative clauses in 28 hours of recorded speech.

A fair number of subjunctive occurrences in relative clauses (n= 54) were found in 20th century HF. If one were to situate this with respect to the total number of subjunctives extracted

(n= 626), subjunctives in this context represent 9% of the total, which falls somewhere between Lagerqvist's (2003) and Lim's (1989) result. Nearly all of the subjunctive occurrences found under relative clauses in 20th century HF were found under the relativizers *qui* (67%, n= 36) and *que* (30%). The raw number of relatives hosting at least one subjunctive according to identity of relativizer in the HF data is given in Table 3.33.

Relativizer	**N**	**%**
Qui	36	67
Que	16	30
Où	1	2
Dont	1	2
TOTAL	54	101[69]

Table 3.33: Total number of subjunctives occurrences under relative clauses found in 20th century HF.

While the analysis thus far has confirmed the presence of the subjunctive in this context, just like the other studies of HF, it still is not clear if this represents a productive use of the subjunctive. In order to address this, all (n= 3184) of the non-applications (i.e., relative clauses headed by *qui* which contained an unambiguous subjunctive, indicative or conditional embedded verb) were extracted. The focus here is solely on relatives headed by *qui* since it selects the largest proportion of subjunctive occurrences compared to the other relativizers[70].

[69] Due to rounding.

[70] A fairly large proportion also appear under *que* (30%) but due to the extremely high frequency of *que* in spoken French (it occurs in many contexts other than relative clauses), the extraction of all of them was beyond the scope of this study (e.g., there are approximately 14,000 occurrences of *que* in 20th century HF alone).

The results presented in Table 3.34 show that the subjunctive in relative clauses introduced by *qui* is actually quite rare, occurring at a rate of about 1% in contemporary HF. This is a far cry from Sand's finding of 78% (n= 14/18)[71].

Rate of subjunctive in relative clauses headed by *qui*	1.1 % (n= 36/3184)
% of relatives preceded by a subjunctive governor	50% (n= 18/36)

Table 3.34: Rate of subjunctive in relative clauses headed by *qui* and % of these relatives preceded by a known subjunctive governor in HF.

In addition, half of the tokens were closely preceded by a known subjunctive governor, as in the examples (17) and (18). This effect was also noted qualitatively by Lim (1989), cited in Blanche-Benveniste et al. (1990:198), and in Sand (2003). Lim refers to it as le "*subjonctif aimanté*" (i.e., the subjunctive attraction phenomenon), an effect that can be attributed to the fact that a subjunctive in a relative clause is "licensed when the noun phrase containing the clause is dependent on an operator which would trigger the subjunctive in a complement clause" (Portner 2011:1285). At a rate of only 1%, with at least half of those preceded by a known subjunctive governor, it is clear that subjunctive use in this context is not at all productive in HF.

(18) « J'aimerais faire un voyage qui ***sorte*** (S) un petit peu de l'ordinaire. » (ESLO/67/306)

'I would like to take a trip that is a bit out of the ordinary.'

[71] Sand (2003) found a much higher rate of subjunctive in relative clauses. It is possible that he only extracted restrictive relative clauses since the examples he provides (n= 2) are of this type, but he does not explicitly state this. In the context of this study, all relative clauses headed by *qui* (restrictive and non-restrictive) were extracted, but even if this had been limited to only restrictive clauses, it would still not account for the extreme difference in the rate of subjunctive in the two varieties (1% > 78%). Additionally, the subjunctive is theoretically possible in unrestrictive relatives (Togeby 1966). All possible contexts in which the subjunctive *could have* occurred are included in this analysis (Principle of Accountability). While this may not be the best way to define the variable context, the validity of the comparison is strengthened by the fact that the extraction and analysis was performed in the same manner and by the same analyst.

Togeby (1966) suggests that the subjunctive only occurs in restrictive relative clauses when the antecedent is indefinite or has no article. Lim (1989, cited in Blanche-Benveniste et al. 1990) also suggests that the antecedent is usually indefinite in the occurrences analyzed in her study, but does not provide quantitative results. She also found that one embedded verb is found more frequently than others in this context: *pouvoir*. In the present study, the same pattern emerged with regards to the definiteness of the antecedent: 81% (n= 29/36) are indefinite. However, *pouvoir* only represents 17% (n= 6/36) of the embedded verbs with subjunctive morphology under *qui*-relatives.

Is the same pattern of use present in QF[72]? In order to answer this question, all of the relative clauses headed by the relativizer *qui* in a subsample of the Ottawa-Hull corpus were extracted (these were not analyzed in Poplack, Lealess & Dion 2013). The results (Table 3.35) reveal that the subjunctive *does* indeed exist in this context in QF, but at the same exceedingly rare rate that it does in HF.

	20th HF	**20th QF**
Rate of subjunctive in relative clauses headed by *qui*	1.1% (n=36/3184)	0.05% (n= 9/1708)
% of relatives preceded by a subjunctive governor	50% (n=18/36)	44% (4/9)

Table 3.35: Rate of subjunctive in relative clauses headed by *qui* in HF and QF.

Here too, nearly half of the occurrences of the subjunctive in relative clauses in QF appear after or in close proximity to a known subjunctive governor. Additionally, most of the antecedents are indefinite (67%, n= 6/9) but none of the tokens contain *pouvoir*.

[72] Only relative clauses are considered in the comparative portion since it is not possible to account for the non-applications in main clauses.

Other non-subordinate subjunctives

In perusing the HF data, it was clear that there were several instances of subjunctive use in contexts with no clear subjunctive governor. Main clauses are often deemed the context in which the subjunctive makes (or made) the strongest semantic contribution (Harris 1978). This use of the subjunctive in HF has also been noted in previous studies. Lim (1989, cited in Blanche-Benveniste et al. 1990) found that 6% of all the subjunctive occurrences consisted of fixed expressions of the type given in (19) and (20). She also noted a few other occurrences of main clause subjunctives but did not provide figures to support this. Lagerqvist (2003) found that 11% of the subjunctive occurrences he extracted consisted of main clause subjunctives, most of which were various types of fixed expressions. Sand (2003) also mentions the use of the subjunctive in main clauses in his study of HF but simply states that there are 16 occurrences and does not provide additional details.

A total of 129 subjunctive occurrences in main clauses were extracted from the HF subsamples used in this study, most of which are fixed expressions (74%), as in (19-22). These results are presented in Table 3.36.

(19) « J'vais gérer des personnes donc euh que j'***sois*** (S) allergique ou pas. » (CFPP/YZ/6823)

'I will manage employees so uh- whether I'm allergic or not.'

(20) « Certains professeurs sont en- enfin que ce ***soit*** (S) en maths ou en sciences sont beaucoup moins à cheval. » (ESLO1/BC709/678)

'Some teachers are we-well, whether it be in math or science, are less strict.'

	Context	N
Fixed expressions	*Que ce soit*	60
	Quel que soit	10
	Quoi que ce soit	5
	Que je réfléchisse	2
	Qu'on dise	2
	Que (N) (V_{subj}) ou …	13
	Que= 'whether', 'let'	4
Other contexts		33
	TOTAL	129

Table 3.36: Raw occurrences of 'other non-governed subjunctives' in 20th century HF.

Examples of one type of context (i.e., *que (N) (V_{subj}) ou …)* is given in (19) and (20), and an example of a main clause subjunctive where *que* is used in a manner similar to 'whether' in English is given in (22).

(21) « Que ça ***plaise*** (S) ou que ça ne ***plaise*** (S) pas c'est autre chose. » (ESLO1/1254/429)

'Whether you like it or not is another story.'

(22) « Si l'enfant s'intéresse pas à ce qu'on lui raconte, bah qu'on lui ***apprenne*** (S) du latin qu'on lui ***apprenne*** (S) des maths qu'on lui ***apprenne*** (S) de l'histoire et de la géographie, il retiendra rien. » (ESLO1/DM95/978)

'If the child is not interested in what we are telling him, whether we teach him Latin or teach him math or teach him history and geography, he will not retain anything.'

The remaining 'other non-governed subjunctive' contexts are mostly comprised of occurrences of subjunctive morphology without any clear direct or proximal subjunctive governor; however, most of them are introduced by *que*, which also plays an important role in subjunctive selection under verbal and non-verbal governors, as shown in Sections 3.4.1.1 and

3.4.1.2. In addition, many of them appear to be speech errors or instances of hypercorrection, without any clear semantic motivation, as in (23) and (24).

(23) « …et donc il est tombé un peu malade là ça ***ait*** (S) un an et demi qu'il est euh deux ans qu'il est en arrêt. » (ESLO2/BC41/389)

'…and so, he got a little sick there, it was a year and a half that he uh two years since he's been on sick leave.'

(24) « …il faut vraiment s'entraîner parce que […] il le ***lise*** (S) qu'une fois. » (ESLO2/KC3/531)

'…you have to really train yourself because he only reads it once. '

Summary

In this section, it was revealed that the use of the subjunctive in both of the non-governed contexts examined here is rare at best. Main clause subjunctives consist mostly of fixed expressions, while subjunctive rates under relative clauses headed by *qui* are extremely low. Comparisons with QF revealed that the same is very much the case there, at least in the case of relative clauses (note that subjunctive use in main clauses was not compared across the two varieties). While it is clear that both HF and QF exhibit nearly identical patterns of subjunctive use in relative clauses headed by *qui*, these results are not easily compared with the other HF studies that have looked at this context. Two of them (Lagerqvist 2003 and Lim 1989) do not include non-applications; therefore, it is impossible to determine the rate of occurrence in this context. Sand (2003) noted a much higher rate of subjunctive in relative clauses than what was found in this study (78% vs. 1%) but how he defined the variable context is not apparent. It is clear, however, that the lack of subjunctive found in these two contexts aligns perfectly with

Harris' (1978) suggestion that it has largely disappeared from the only contexts that were unequivocally semantically motivated at a much earlier stage.

3.4.1.5 Discussion

In previous sections, subjunctive use in contemporary HF was analyzed using the variationist method and the results were subsequently compared with those of Poplack, Lealess & Dion (2013) for QF in order to address two of the research questions which form the basis of this study (the diachronic component will be discussed in the next section):

1) Is the subjunctive used to the same degree in HF as in QF?

2) Are there parallels in the linguistic conditioning of the variability in both varieties?

Through the use of this systematic and accountable method, a detailed picture of how and where the subjunctive occurs in 20th century HF has emerged. Moreover, this method allowed for several methodological shortcomings, apparent in other HF subjunctive literature, to be overcome, such as the heavy reliance on prescriptivist accounts of usage as a baseline or the dependence on rates of occurrences alone without considering any other conditioning factors (or even non-applications in some cases). Had only rates of occurrence been considered, it would have led to an entirely different set of conclusions.

Several important findings emerged from this comparative analysis. First, subjunctive use in HF exhibits limited productivity. Recall that subjunctive productivity was defined at the outset as freely occurring without lexical or structural restrictions in all eligible contexts, and as contributing semantic value to the utterance. With the help of Poplack's three-way measure, a distinct pattern emerged, much like the one found by Poplack, Lealess & Dion for QF: only a handful of verbal and non-verbal governors exhibit high rates of co-occurrence with the

subjunctive, while simultaneously accounting for a large proportion of both the governor and subjunctive morphology pools. In addition to clear lexical effects, morphosyntactic factors also condition subjunctive use (the presence of *que*, the absence of intervening material between the governor and the embedded verb, and a non-suppletive embedded verb), not semantic ones. Where a semantic factor group at first appeared to be significant, further analysis revealed underlying lexical effects. Moreover, the subjunctive is rare in both relative clauses (where its use is usually primed by a nearby subjunctive governor) and in non-governed clauses (where it mostly occurs in fixed expressions).

Previous studies of HF (outlined in Section 3.2.2) did not provide a clear picture of subjunctive use due to various methodological disparities with regard to the type of data, the definition of the variable context, and the semantic classification of the subjunctive governors. Additionally, rates of occurrence (when provided) vary widely from one study to the next, some higher (e.g., Sand 2003) and some lower (e.g., Chauveau 1998) than what was found here. Some of the findings that emerged from the present study have been noted or alluded to in the previous work on the HF subjunctive; however, the extent to which the subjunctive is lexically and structurally (and not semantically) restricted in contemporary HF has not been identified until now. For example, *pour que* and *falloir que* have both been previously recognized as having strong and frequent associations with the subjunctive (e.g., Gougenheim et al. 1964), but the extent of their proportional dominance was never noted. Sand (1983) suggested that distance between a governor and the embedded verb may have explanatory value but never quantitatively tests this hypothesis, while Nordahl (1969) found that distance between the governor and the embedded verb was only applicable to what he defined as subjective verbs. Sand (2003) was the

only one[73] to mention the identity of the embedded verb, stating that about two thirds of embedded subjunctives are made up of *être, avoir, faire, and aller* but does not push the analysis beyond this observation to discover that, while these embedded verbs may the most frequent, the rate of subjunctive with two of them (*être, avoir*), is actually quite a bit lower than with the others.

With regards to semantic class, the results of this analysis jibe at first glance with what has previously been reported for HF: subjunctive use is (near) categorical under volition and emotion verbal governors. However, the analysis of the lexical items contained within the classes led to the conclusion that this supposed semantic effect was masking a lexical one. This study is not the first to suggest that subjunctive use in HF is not semantically conditioned (e.g., Gross 1978) but the use of the variationist method allowed for subjunctive productivity to be assessed and the fine details of subjunctive use to be revealed through analysis of lexical, structural, and semantic factors simultaneously.

Another important finding to emerge from this study was that both HF and QF exhibit many striking similarities when it comes to proportional disparities across governors, their associated rates of subjunctive, as well their underlying grammar. In both varieties, there is one main verbal (*falloir*) and non-verbal (*pour que*) governor that stands apart from the rest in that they exhibit (near) categorical subjunctive rates, while accounting for large proportions of both the governor and subjunctive morphology pools.

[73] To my knowledge.

VERBAL	**20th century HF**			**20th century QF**		
	% pool	Rate subj.	% subj.	% pool	Rate subj.	% subj.
Importance of top-tier governors	53	92	67	80	89	92
Number of top-tier governors	5			3		
Identity of most important top-tier governor	*falloir que*			*falloir que*		
Identity of other top-tier governors	*aimer que*, *vouloir que*, *croire que* (neg.), *penser que* (neg.)			*aimer que*, *vouloir que*		
Morphosyntactic conditioning factors	Absence of intervening material Presence of *que*			Absence of intervening material Presence of *que* Suppletive embedded verbs		
NON-VERBAL	**20th century HF**			**20th century QF**		
	% pool	Rate subj.	% subj.	% pool	Rate subj.	% subj.
Importance of top-tier governors	49	96	64	84	85	88
Number of top-tier governors	5			5		
Identity of most important top-tier governor	*pour que*			*pour que*		
Identity of other top-tier governors	*avant que*, *à moins que*, *sans que*, *bien que*			*avant que*, *mais que*, *soit que*, *jusqu'à tant que*		
Morphosyntactic conditioning factors	Presence of *que* Regular embedded verb			Presence of *que*		

Table 3.37: Summary of findings in HF and QF with regards to disproportionate distributions of subjunctive governors and internal conditioning.

In both varieties, there is also a set of influential governors whose proportional role outweighs that of the other governors and, while their number and lexical identity varies slightly, it is clear that subjunctive use is largely lexically restricted in both cases. Moreover, structural factors predict subjunctive use over semantic ones in both HF and QF with the presence of *que* and the absence of intervening material both playing a role in most of the data sets tested.

While the results of this analysis have shown that subjunctive productivity is limited in both varieties, there is evidence to suggest that it is more restricted in QF than in HF. The top-tier verbal and non-verbal governors account for a much larger proportion of both the subjunctive morphology and governor pools in QF than they do in HF (Table 3.37). Despite that the rate of subjunctive being higher with these governors in HF, the fact that they account for less of the data overall means that subjunctive use is spread out over a larger number of remaining contexts. These remaining contexts are more frequent in HF, as evidenced by the greater number of governors found to co-occur with the subjunctive, but most are exceedingly infrequent.

The variationist method provided the tools required to paint a clear picture of subjunctive use in contemporary HF. It is not as robust as many assume and despite an overall rate of 71%, its use is largely restricted to a small number of governors and is practically non-existent in relative clauses or main clauses. But are these patterns the result of change? Poplack found evidence of entrenchment of these patterns of use across time in QF, but will a similar pattern of entrenchment emerge from the analysis of HF? These are the questions that will be explored in the next section, where the findings presented here are compared to 21st century QF and HF.

3.4.2 Diachronic Analysis of subjunctive use in HF and QF

A synchronic (horizontal) comparative analysis of subjunctive use in 20th century HF and QF was presented in the last section. In this section, the focus will be on if and how subjunctive use has changed over time in both varieties. The motivation for undertaking a diachronic analysis is twofold. First, there are conflicting claims regarding the decline of the subjunctive in contemporary HF (e.g., Dreer 2007; Lagerqvist 2003) and second, Poplack, Lealess & Dion (2013) found strong evidence of an increased association (i.e., entrenchment) between the subjunctive and certain governors over time in their diachronic analysis of QF. While the results of the present study (presented in the first part of this chapter) showed that subjunctive use is strongly associated with a handful of governors in 20th century HF as well (albeit to a lesser degree than in QF), it is not clear if this effect is increasing over time and thus resulting in a similar pattern of entrenchment. A diachronic analysis between 20th and 21st century HF will identify if the entrenchment trajectory that is present in QF is also present in HF and if the change is moving at a rate that can be captured in this timeframe.

In order to complete this analysis, two 21st century HF corpora (*Corpus du Français Parlé Parisien des années 2000 and Enquêtes Sociolinguistiques à Orléans 2008-2010)* are used in addition to the 20th century one (*Enquêtes Sociolinguistiques à Orléans 1968-1970)* that has been the source of HF data up until this point. As mentioned previously, the two 21st century corpora are combined since preliminary analyses revealed no significant difference between the two when it comes to rate of subjunctive use or level of formality.

In Section 2.3.1, two different ways in which an analyst can measure language change (*real time* and *apparent time*) were discussed in detail. A linguistic analysis in *real time* refers to the comparison of the linguistic patterns of speakers at two different points in time. The group of

speakers can be the same or different. A linguistic analysis in *apparent time* refers to the study of the linguistic patterns of different age groups at one single point in time. Differences between groups can be indicative of change in progress, under the assumption that the speech of the oldest speakers represents the state of the language at the time they were acquiring it, a point in time prior to the one represented by the youngest speakers (e.g., Labov 1994). An apparent time analysis takes for granted the assumption that one's linguistic patterns remain stable over the course of one's lifetime, and is a methodological tool commonly used in sociolinguistic research in order to assess change in the absence of real time data (Walker 2010:96–97). While a diachronic component has been built into this study, given its short time span, it will be useful to appeal to the apparent time angle as well where appropriate.

The 20th century HF corpus was recorded between 1968 and 1970 and the 21st century corpora were recorded between 2005 and 2011, producing a real time span of approximately 35-40 years in which to identify any potential changes in subjunctive use. Within each corpus, a variety of age groups are represented which will also allow for the use of the apparent time construct when needed. However, the majority of the analysis will focus on the real time component. While it may be the case that this time span is too limited to reveal change in progress in the use of the subjunctive given that linguistic change tends to occur gradually over long periods of time, the possibility still exists. In addition, Poplack, Lealess & Dion (2013) found some evidence of change across time within a similar time frame although the majority of their evidence for entrenchment of the subjunctive with certain governors and embedded verbs came from their comparison between the 19th and 20th century (a point of comparison not available here).

3.4.2.1 Verbal governors

Overall rate

Recall that the rate of subjunctive under verbal governors in 20th century HF is 71% (n= 520/732). When compared to 21st century HF, it is clear that the overall rate (70%, n=298/423) has remained virtually unchanged. In QF, there is little change in the overall rate from one time period to the next, but subjunctive use in this context remains higher than HF in both (20th: 76%, n= 1953/2569; 21st: 79%, n= 663/841) (Figure 3.4).

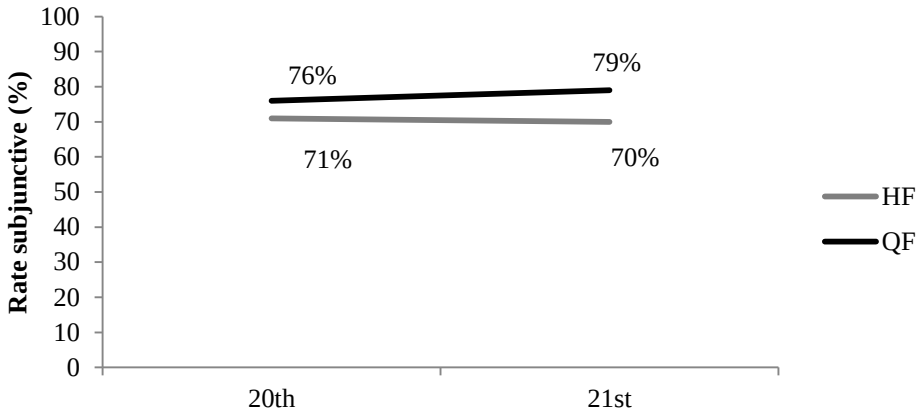


Figure 3.4: Overall rate of subjunctive under verbal governors in 20th and 21st century HF and QF[74].

Lexical identity of governors

Despite finding subjunctive morphology under 49 verbal governors in 20th century HF, more than two thirds of it was actually accounted for by only five of them (*falloir, vouloir, aimer, penser* (neg.), *and croire* (neg.)) of which *falloir* was by far the most important, accounting alone for 42% of all of the subjunctive morphology. In 21st century HF, roughly the

[74] Results for QF adapted from Poplack, Lealess & Dion (2013: 166).

same number of verbal governors (n= 52)[75] co-occur with the subjunctive at least once, but as was the case in the 20th century, the rate of associations with the subjunctive varies widely (Table 3.38)[76].

Some interesting patterns emerge from this diachronic comparison. First, nearly half of the governors that do persist across the two time periods are fairly frequent (n= 11/24), yet the majority of these (n= 8/11) exhibit a decrease in their rate of co-occurrence with the subjunctive (in some cases quite minimally). The two most frequent governors (*falloir* and *être +attribute*) remain stable, with the former's associations with the subjunctive strengthening, which is indicative of entrenchment. Second, the majority of the governors that are idiosyncratic to the 20th century are infrequent (n= 19/25). Third, every single governor idiosyncratic to the 21st century is infrequent (n= 28/28) and many (n= 21/28) are nonce tokens. Of the twenty nonce governors new to the 21st century, the majority occur in a structural context favourable to the subjunctive (with *que:* 100%; with no parenthetical intervening material: 76%). The emergence of *trouver* +adj. should also be noted, which is structurally similar to the already established *être* +attribute, already strongly linked to the subjunctive in the 20th century. In fact, nearly half of the governors (n= 13/28) that are idiosyncratic to the 21st century can be accounted for by one of the following constructions: *trouver x que* or *être/avoir x que.* In addition, many of the new contexts in which the subjunctive is found in the 21st century are not those in which one would typically expect to find a subjunctive according to prescriptive norms (e.g., *faire rigoler*).

[75] Of the 52 verbal governors, the construction *être X (adj/noun) que* is counted as a single governor despite the fact that eighteen different lexical items appear in this context: *bien* 'good', *content* 'happy', *dégoutânt* 'disgusting', *dommage* 'unfortunate'*, étonné* 'surprised', *extraordinaire* 'extraordinary', *hors question* 'out of the question', *inconceivable* 'inconceivable'*, l'idée* 'the idea'*, marrant* 'funny'*, impossible* 'impossible', *possible* 'possible', *prioiritaire* 'having priority', *pas sûr* 'not sure', *rare* 'rare', *regrettable* 'regrettable*, une bonne chose* 'a good thing', *voulu* 'wanted'.

[76] The translation for each of the verbal governors found in HF is given in Appendix A.

To summarize, there is a lot of fluctuation across the two time periods with regards to the lexical identities of the verbal subjunctive governors; however, it is mostly found amongst those that occur infrequently in the data. Such fluctuations and the emergence of novel governors in the structural context that is favourable to the subjunctive represents yet another argument against the semantic hypothesis and thus, against the idea that the subjunctive is productive in contemporary HF.

	Persistent governors						Governors idiosyncratic to 20th century				Governors idiosyncratic to 21st century			
	20th			21st				20th				21st		
	% subj	N subj	Tot N	% subj	N subj	Tot N		% subj	N subj	Tot N		% subj	N subj	Tot N
falloir	98	216	221	95	120	126	*voir*	7	3	44	*espérer*	14	1	7
être +attribute	84	49	58	90	28	32	*sembler*	9	2	22	*se souvenir*	100	3	3
vouloir	96	54	56	90	27	30	*admettre*	52	11	21	*avoir besoin*	100	3	3
penser (neg.)	72	36	50	58	14	24	*vouloir dire*	6	1	16	*connaître*	33	1	3
croire (neg.)	72	26	36	40	2	5	*faire*	17	2	12	*avoir marre*	100	2	2
pouvoir dire (neg.)	42	10	24	39	7	18	*valoir mieux*	100	10	10	*avoir le sentiment* (neg.)	100	2	2
aimer (mieux/bien)	95	18	19	79	11	14	*supposer*	29	2	7	*avoir réussi à ce*	100	2	2
c'est pas	22	4	18	17	1	6	*faire attention*	14	1	7	*accepter* (neg.)	100	1	1
ce qui fait	6	1	18	12	3	25	*dire*	20	1	5	*accorder*	100	1	1
comprendre	50	7	14	27	4	15	*avoir pas de raisons*	100	3	3	*avoir aucun souvenir*	100	1	1
préférer	100	10	10	100	4	4	*éviter*	100	3	3	*déplaire*	100	1	1
attendre	100	6	6	100	3	3	*se pouvoir*	67	2	3	*faire rigoler*	100	1	1
trouver (neg.)	50	3	6	67	4	6	*craindre*	100	2	2	*gêner* (neg.)	100	1	1
suffire	80	4	5	75	3	4	*faire changement à ce*	100	2	2	*mériter*	100	1	1
demander	100	3	3	100	1	1	*regretter*	100	2	2	*promettre*	100	1	1
tenir à	100	3	3	100	3	3	*surprendre*	100	2	2	*rester*	100	1	1
faire en sorte	100	3	3	100	4	4	*permettre*	50	1	2	*supporter* (neg.)	100	1	1
avoir/être peur	100	2	2	80	8	10	*faire du mal*	50	1	2	*trouver bien*	100	1	1
souhaiter	100	2	2	100	3	3	*s'étonner*	100	1	1	*trouver chouettte*	100	1	1
avoir (x) chance	50	1	2	75	3	4	*avoir à coeur*	100	1	1	*trouver dommage*	100	1	1
arriver	100	1	1	25	2	8	*exiger*	100	1	1	*trouver dramatique*	100	1	1
avoir l'impression	100	1	1	18	6	33	*faire plaisir*	100	1	1	*trouver insalubre*	100	1	1
arranger	100	1	1	100	1	1	*paraître impensable*	100	1	1	*trouver pas mal*	100	1	1
envisager	100	1	1	100	1	1	*rendre*	100	1	1	*valoir le coup*	100	1	1
							trouver logique	100	1	1	*plaire*	100	1	1
											avoir envie (neg.)	100	1	1
											être/avoir moyen (neg.)	100	1	1
											considérer (neg.)	100	1	1

Table 3.38: Lexical identity of verbal governors in 20th and 21st century HF according to frequency of governor (descending order).

Further analysis of the top verbal governors identified in 20th century HF reveals that *falloir* neither increases nor decreases in terms of its productivity across the time period examined. Similarly, Poplack, Lealess & Dion (2013) found little change with *falloir* across roughly the same time period, with the exception of a modest increase in rate of subjunctive under it, which rose from 89% to 94% (Table 3.39). In both periods, *falloir* continues to occupy a large proportion of the governor and subjunctive morphology pools, all while maintaining its strong association with the subjunctive. However, as was the case in the synchronic comparison (Section 3.4.1), the two varieties continue to exhibit proportional discrepancies with regards to *falloir*: it remains much more prominent relative to the other verbal governors in QF than in HF.

	20th century HF				**21st century HF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
falloir que	30	98	216/221	42	30	95	120/126	40
others	70	59	304/511	58	70	60	178/297	60
Total	100	71	520/732	100	100	70	298/423	100
	20th century QF				**21st century QF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
falloir que	65	89	1493/1669	76	63	94	498/530	75
other	35	51	460/900	24	37	53	165/311	25
Total	100	76	1953/2569	100	100	79	663/841	100

Table 3.39: Importance of *falloir que* vs. other verbal governors as shown by its portion of the verbal governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th and 21st century HF and QF[77].

With regards to the other verbal governors, the same five (with the exception of *croire que* (neg.)) continue to account for the largest proportion of subjunctive morphology all while maintaining high rates of co-occurrence with the subjunctive in HF (Table 3.40). Once again, there is little difference in each of these verbal governors across the time span examined here in

[77] Results for QF adapted from Poplack, Lealess & Dion (2013:166).

terms of their respective proportions of both the governor and subjunctive morphology pools; however, their associations with the subjunctive have decreased. The only notable exception is *croire que* (neg.), which has dropped off both in terms of rate and portion of the subjunctive morphology pool in 21st century HF.

	20th century HF				**21st century HF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
		%	**n**			**%**	**n**	
falloir	30	98	216/221	42	30	95	120/126	40
vouloir	8	96	54/56	10	7	90	27/30	9
aimer	3	95	18/19	3	3	79	11/14	4
penser (neg.)	7	72	36/50	7	6	58	14/24	5
croire (neg.)	5	72	26/36	5	1	40	2/5	1
Total (top-tier verbal governors)	53	92	350/382	67	47	87	174/199	59
	20th century QF				**21st century QF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
		%	**n**			**%**	**n**	
falloir	65	89	1493/1669	76	63	94	498/530	75
vouloir	11	91	249/273	13	9	93	71/76	11
aimer	4	67	66/99	3	2	92	12/13	2
Total (top-tier verbal governors)	80	89	1808/2041	92	76	94	581/619	88

Table 3.40: Importance of top-tier verbal governors as shown by each of the verbal governor's portion of the verbal governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in all four cohorts[78].

In their analysis of 21st century QF, Poplack, Lealess & Dion (2013) found that the same three verbal governors maintained their disproportionate status relative to the other governors that was found in the 20th century. However, the overall rate of these three combined did rise slightly due to an increase in the overall rate with each of them (Table 3.40). The entrenchment

[78] QF results adapted from Poplack, Lealess & Dion (2013:166).

of the subjunctive with the top verbal governors in QF continues to grow, while in HF, it has remained fairly stable, with the exception of *croire que* (neg.)[79].

Summary

While Poplack, Lealess & Dion identified a loss of productivity coupled with the strengthening of entrenchment of the subjunctive in certain contexts between 20th and 21st century QF, the diachronic analysis presented here did not reveal any significant change in HF during approximately the same time span. Despite this, the relative stability of the highly frequent governors and the fluctuations of the rarer ones speak to a similar pattern of entrenchment as what was found in QF. In other words, subjunctive productivity was limited in 20th century HF and continues to be in the 21st century data, but the effect does not appear to be strengthening.

3.4.2.2 Non-verbal governors

Overall rate

In terms of the overall rate of subjunctive use under non-verbal governors in the 20th century, the analysis of the HF data revealed a lower rate (73%) than what Poplack, Lealess & Dion (2013) found in QF (82%) in the same time period. The diachronic analysis of subjunctive use under non-verbal governors across time in HF shows that, while the rate decreases from 73% (n= 106/145) to 68% (n= 100/147) between the 20th and 21st century, this difference is not statistically significant (p= .3702), according to Fisher's exact test. Poplack, Lealess & Dion (2013) actually found a decrease in overall usage between the 20th (82%, n= 163/198) and the

[79] Since there is no evidence to suggest change in subjunctive use in HF across the 20th and 21st century, internal conditioning is not examined here. For an overview of the internal conditioning for verbal and non-verbal governors across this period of time in QF, see Poplack, Lealess & Dion (2013).

21st century (66%, n= 164/249) but attribute this to a disproportion of high frequency, low subjunctive governors in the 21st century data (Figure 3.5).

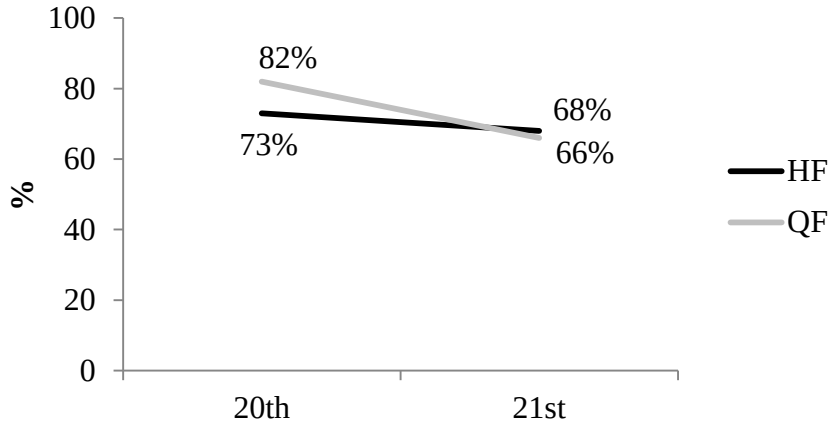


Figure 3.5: Overall rate of subjunctive under non-verbal governors in 20th and 21st century HF and QF[80].

Lexical identity of governors

In Section 3.4.1.2, it was shown that subjunctive use is concentrated with a handful of non-verbal governors in both 20th century varieties. In this section, the stability of this trend across time will be explored.

Subjunctive morphology is found under a total of 19 non-verbal governors in 21st century HF (down from 23), most of which (n= 15) persist across the two time periods (Table 3.42)[81]. The most frequent persist across time and generally remain stable in terms of their rate of occurrence with the subjunctive. All of the idiosyncratic governors are relatively infrequent, all occurring four times or less in the data in a potential subjunctive-selecting context.

[80] Results for QF adapted from Poplack, Lealess & Dion (2013: 176).
[81] The translation for each of the non-verbal governors found in HF is given in Appendix A.

	Persistent governors						Governors idiosyncratic to 20th century				Governors idiosyncratic to 21st century			
	20th			21st				20th				21st		
	% subj	N subj	Tot N	% subj	N subj	Tot N		% subj	N subj	Tot N		% subj	N subj	Tot N
pour que	100	38	38	100	38	38	*au/en cas que*	100	1	1	*au lieu que*	100	2	2
même que	71	12	17	50	1	2	*façon à ce que*	100	2	2	*à part que*	50	1	2
le fait que	13	2	15	26	11	43	*à condition que*	100	2	2	*quoi que*	75	3	4
à moins que	85	11	13	100	1	1	*plutôt que*	50	1	2				
avant (même) que	100	9	9	80	16	20	*sur le principe que*	50	1	2				
bien que	86	6	7	67	2	3	*après que*	67	2	3				
maintenant que	14	1	7	20	1	5								
malgré que	60	3	5	57	4	7								
sans que	100	4	4	100	12	12								
des fois que	25	1	4	100	1	1								
le temps que	100	2	2	100	1	1								
de manière à ce que	100	2	2	100	1	1								
non pas que	50	1	2	100	1	1								
jusqu'à tant que	100	1	1	100	3	3								
jusqu'à ce que	100	1	1	100	1	1								

Table 3.41: Lexical identity of non-verbal governors in 20th and 21st century HF.

Do the top non-verbal governors, as identified in Section 3.4.1.2, also pattern similarly across time? Starting with HF, the results presented in Table 3.42 show first, that *pour que* continues to occupy the top spot in both 21st century HF cohorts in terms of proportional importance, while its categorical association with the subjunctive persists. Second, the combination of these same five non-verbal governors continues to be proportionally over-represented in relation to the others, while their combined rate of co-occurrence with the subjunctive remains near categorical. All of these factors point to the persistence of limited subjunctive productivity under non-verbal governors in HF. However, there are slight changes within the group of top non-verbal governors, in particular that the proportional importance of *à moins que* and *bien que* has decreased, while it has increased for *avant que* and *sans que.*

	20th century HF				**21st century HF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
		%	**n**			**%**	**n**	
GOVERNOR								
pour que	26	100	38/38	36	26	100	38/38	38
à moins que	9	85	11/13	10	1	100	1/1	1
bien que	5	86	6/7	6	2	67	2/3	2
avant que	6	100	9/9	8	14	80	16/20	16
sans que	3	100	4/4	4	8	100	12/12	12
Total (top-tier non-verbal governors)	49	96	68/71	64	51	93	69/74	69

Table 3.42: Importance of top-tier vs. other non-verbal governors as shown by each of the non-verbal governor's portion of the governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in 20th and 21st century HF**.**

In QF (Table 3.43), the governors forming part of the top-tier class fluctuate across time with the exception of *avant que* and *pour que*. Most appear to be losing importance (i.e., *mais que, jusqu'à tant que*) at the expense of *pour que,* whose portion of the subjunctive morphology pool rises drastically across this period of time (33% > 57%) all while maintaining a near

categorical rate of co-occurrence with the subjunctive. Poplack, Lealess & Dion (2013) attribute this to the fact that less frequent governors are more volatile, while the frequent governors tend to remain stable, an explanation that can also be adopted for HF. The most frequent non-verbal governor, *pour que*, retains its top position with respect to the others in both time periods and in both varieties, with its position strengthening in QF.

	20th century QF				**21st century QF**			
	% Pool	**Rate Subj.**		**% Subj.**	**% Pool**	**Rate Subj.**		**% Subj.**
		%	**n**			**%**	**n**	
GOVERNOR								
pour que	27	98	53/54	33	39	95	93/98	57
avant que	16	87	27/31	17	4	80	8/10	5
mais que	18	91	32/35	20	/	/	/	/
jusqu'à tant que	9	82	14/17	9	/	/	/	/
soit que	14	50	14/28	9	/	/	/	/
à moins que	/	/	/	/	5	77	10/13	6
après que	/	/	/	/	8	40	8/20	5
Total (top-tier non-verbal governors	84	85	140/165	88	56	84	119/141	73

Table 3.43: Importance of top-tier vs. other non-verbal governors as shown by each of the non-verbal governor's portion of the governor pool (% pool), the portion of subjunctive morphology found under each (% subj.) and the overall rate of subjunctive under each (Rate subj.) in all four cohorts.

Summary

There is no clear evidence of change across 20th and 21st century HF when it comes to subjunctive use under non-verbal governors unlike in QF, where Poplack, Lealess & Dion (2013) noted strengthening of the subjunctive use in certain contexts, namely with *pour que*, which is indicative of increased entrenchment amongst non-verbal governors.

3.4.2.3 Embedded verbs

The diachronic analysis of subjunctive use in HF was extended to the identity of the embedded verb (Table 3.44), but revealed virtually no difference across time. In other words, there is no evidence to suggest that HF is following the same trajectory of entrenchment as QF with regards to the lexical identity of the embedded verb, at least within the time timeframe examined here. It is worth noting that the main evidence supporting the QF's trajectory of entrenchment with specific embedded verbs is found between the 19th and 20th century (see Poplack, Lealess & Dion 2013). They also noted few differences with regards to the embedded verb between the 20th and 21st century.

	20th				21st			
	% Pool	Rate Subj. %	N	% Subj.	% Pool	Rate Subj. %	N	% Subj.
Total top-tier embedded verbs (*être, avoir, aller, faire*)	73	69%	441/637	70	73	66%	275/414	69
Other embedded verbs	27	77%	185/240	30	27	79%	123/156	31

Table 3.44[82]: Importance of the lexical identity of embedded verbs under verbal and non-verbal governors as shown by each of the embedded verb's portion of the pool (% pool), the portion of subjunctive morphology found with each (% subj.) and the overall rate of subjunctive with each (Rate subj.) in 20th and 21st century HF.

In the HF literature, there are claims regarding the use of non-subjunctive embedded verbs out of necessity (i.e., to express temporal reference which cannot be expressed by the subjunctive forms currently used in the language) or as a replacement strategy, leading to the overall decline of the subjunctive. One such strategy involves using other tenses in the embedded clause, which according to some, the subjunctive is not capable of expressing (i.e., the future, the

[82] Both the verbal and non-verbal governors are combined for the analysis of the embedded verb since most important players are common to both data sets.

imperfect, the conditional, or the *passé simple*) (e.g., Lagerqvist 2003; Togeby 1974). In testing this claim, no evidence of such change emerged, with these supposed subjunctive-alternative tenses occupying a combined total of 26% and 25% of the non-subjunctive embedded verbs across the two time periods (Table 3.45).

	20th HF	**21st HF**
Conditional	7% (n=18/251)	2% (n=3/172)
Future	6% (n=14/251)	4% (n=7/172)
Imperfect	13% (n=32/251)	19% (n=32/172)
Other	75% (n=187/251)	76% (n=130/172)

Table 3.45: Tense of embedded verb across time in HF.

3.4.2.4 Non-governed subjunctives

In terms of subjunctive use in *qui*-relative clauses, the results in Section 3.4.1.4 revealed that it is extremely rare in both 20th century HF and QF, with a subjunctive rate of 1% or less in both varieties. A similar result emerges in 21st century HF, however, the rate actually decreases slightly from 1.1% to 0.5% across this time period. The portion of these tokens that were preceded by a known subjunctive governor remains perfectly stable at 50% (Table 3.46).

	20th HF	**21st HF**
Rate of subjunctive in relative clauses headed by *qui*	1.1 % (n= 36/3184)	0.5% (n= 14/2765)
% of relatives preceded by a known subjunctive governor	50% (n= 18/36)	50% (n= 7/14)

Table 3.46: Rate of subjunctive in relative clauses headed by *qui* and % of these relatives preceded by a known subjunctive governor in 20th and 21st century HF.

Recall that the use of the subjunctive in other non-subordinate clauses in 20th century HF was largely (74%, n= 96/129) limited to fixed expressions. Over time, the proportion of such

expressions increases slightly to 80% (n= 20/102) in the 21st century (Figure 3.6), suggesting a further decrease in the already severely limited productivity of the subjunctive in this context.

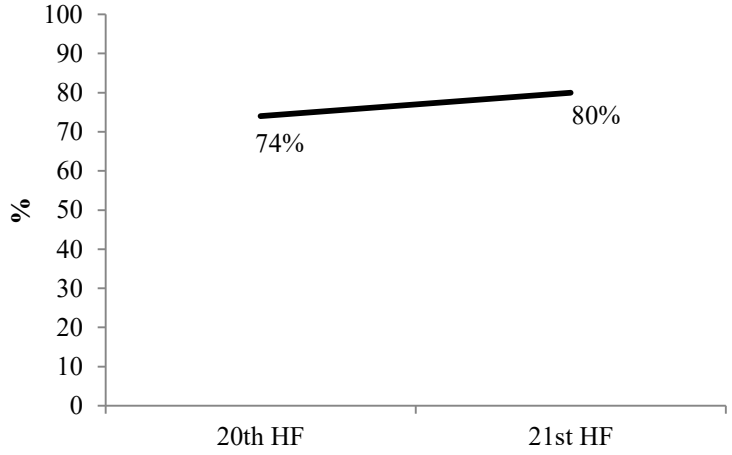


Figure 3.6: Proportion of subjunctives in non-subordinate contexts that are comprised of various fixed expressions.

3.4.2.5 External (social) factors

Even though change in the use of the subjunctive in HF was not confirmed in real time, it is still possible to find such evidence in an apparent time analysis. However, as is shown below, the findings of the preceding sections are bolstered by the analysis of social factors. There is no evidence of change in apparent time either (Figure 3.7).

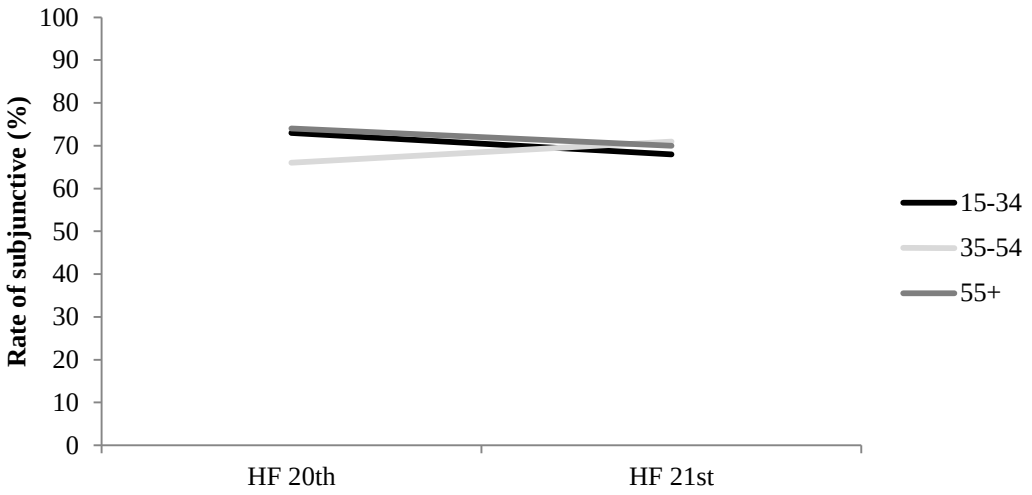


Figure 3.7: Overall rate of subjunctive under verbal and non-verbal governors combined in 20th and 21st century HF according to age.

A difference in subjunctive use according to speaker sex could also be indicative of change in progress. However, there are no clear (i.e., statistically significant) differences between the cohorts across time (Figure 3.8).

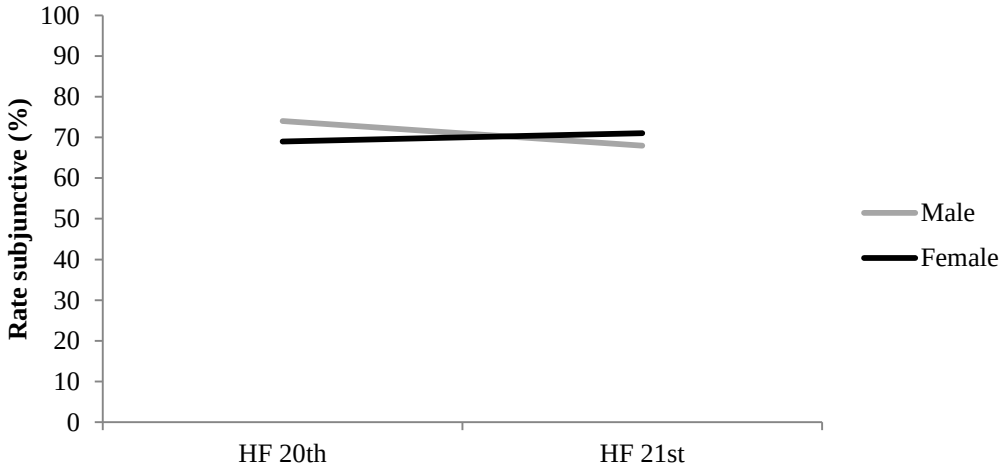


Figure 3.8: Overall rate of subjunctive under verbal and non-verbal governors combined in 20th and 21st century HF according to sex.

These findings further confirm the results presented throughout this chapter in that subjunctive use in HF is stable and not undergoing any significant change, at least in the period of time under analysis here.

3.4.3 Discussion of diachronic analysis

In this section, subjunctive use across two HF time periods (the 20th and 21st centuries) was analyzed in order to identify possible change. The results for HF were subsequently compared with those of Poplack, Lealess & Dion (2013) for roughly the same two time periods in QF. With the both the verbal and non-verbal governors, virtually no change was identified in HF. Instead, remarkable stability was found in the rate of subjunctive, with virtually the same few governors occupying the top spots across time. The only major exception was *croire* (neg.), which dropped out of the top-tier verbal governors but this was already the least important of the top governors to begin with. More generally, it appears that the more frequent governors persisted across time, while the non-frequent ones were subject to frequent fluctuations. This is further evidence of the lack of semantic motivation for subjunctive use because if it were, in fact, playing a role, such glaring frequency effects and proportional irregularities would not have been found.

As was suggested earlier, the time span covered in the diachronic analysis may not be sufficient to identify change in progress since linguistic change tends to be a very gradual progress. However, Poplack, Lealess & Dion (2013) identified evidence of entrenchment in roughly the same span in the QF data; therefore, it was important to test whether the same patterns were present in HF. Additionally, this analysis tested some of the claims pertaining to the decline of the subjunctive in previous (mostly synchronic or descriptive) studies of HF (e.g.,

van der Molen 1923; Dreer 2007; Bauche 1929), but found little evidence of a decline within the two time periods examined here.

Another claim in the HF literature was refuted in this study: the replacement strategy of using alternate tenses (e.g., future, conditional, imperfect) or infinitives instead of the subjunctive in the embedded clause (e.g., Lagerqvist 2003; Bauche 1929; Togeby 1966; 1974). No evidence of such change was identified, with embedded verbs conjugated in the future, conditional, or imperfect only representing one quarter of the non-subjunctive embedded verbs and remaining stable across time (25% > 24%). The use of an infinitival complement as a replacement strategy was not tested due to the fact that it is only possible in a limited number of contexts (including *falloir*, which will be examined in more detail in the next chapter).

Poplack, Lealess & Dion's (2013) study revealed that the main evidence of entrenchment of the subjunctive with a handful of governors and embedded verbs in QF originated from the comparison of the 19th century with the 20th and 21st centuries. Since HF appears to exhibit very similar, but less strong, patterns of structural and lexical restriction in subjunctive use, it is possible that it may resemble an earlier stage of QF (i.e., the 19th century). Table 3.47 summarizes some of the main findings in 19th century QF compared with 20th century HF and 20th century QF.

Based on the results presented in Table 3.47, it can be concluded that HF does, in fact, resemble an earlier stage of QF when it comes to the rate of co-occurrence of the subjunctive with *falloir* and the relative number of subjunctive governors. In terms of the overall rate of subjunctive and the proportion of the subjunctive morphology pool occupied by *pour que* and *falloir*, 20th century HF more closely resembles 20th century QF.

	19th century QF[83]	**20th century HF**	**20th century QF**
Overall rate of subjunctive (verbal and non-verbal combined)	57%	71%	76%
falloir (rate subjunctive)	62%	98%	89%
falloir (% subjunctive morphology pool)	62%	42%	76%
pour que (rate subjunctive)	71%	100%	98%
pour que (% subjunctive morphology pool)	53%	36%	33%
Number of verbal governors /1000 words	4.2	5.9	1.3
Number of non-verbal governors /1000 words	2.7	2.2	0.5

Table 3.47: Summary of main finding with regards to importance of top-tier governors in 19th and 20th century QF and 20th century HF (shading indicates similarities across cohorts).

It is known that subjunctive use in QF has entrenched over time to a few restricted contexts and, while similar patterns of use in contemporary HF were identified, there is no evidence to suggest that it has changed as well, at least not in the time period examined here. However, if one were to consider a hypothetical earlier stage where the subjunctive was fully productive, it can be inferred that change toward entrenchment took place. Without an earlier HF benchmark, one cannot say with certainty that change has taken place, but the evidence presented in this chapter suggests that it has.

3.5 Conclusion

This analysis of subjunctive use in HF and subsequent comparisons with subjunctive use in QF revealed some interesting trends. While the subjunctive occurs at an overall lower rate in HF than QF, these differences are actually quite superficial. The underlying grammars are quite similar, particularly with regards to the underlying lexical effects which have been found to

[83] Results for QF adapted from Poplack, Lealess & Dion (2013) and personal communication.

operate in several varieties of Canadian French, namely the proportional importance of a small number of subjunctive governors which all have strong associations with the subjunctive. Unfortunately, the diachronic analysis of subjunctive use in 20th and 21st century HF was too short of a time gap to capture the change that led to its current state. In addition to lexical effects, strong structural effects are also at play, including the presence of *que* and the distance between the governor and embedded verb.

Methodological disparities amongst previous quantitative HF studies of the subjunctive make it hard to proceed with direct comparisons. While a handful of these studies touched upon some of the effects found here, none were able to identify the extent to which lexical and structural factors account for its use in contemporary French, a finding that was enabled through the use of the variationist method. As a result, a more nuanced and complete explanation of its use has emerged, instead of only bits and pieces of the story.

Within the context of this framework, potential semantic effects on subjunctive use were empirically tested, which led to the finding that those that appeared to be operative were, in fact, due to underlying lexical effects. This methodological approach differs immensely from those used in previous HF studies in that, here, the focus is on all of the contexts in which the subjunctive was actually used rather than those in which it is prescribed.

Despite the common belief that the subjunctive is more frequent and more productive in HF, the results of this study show that its productivity is actually quite limited, albeit to a lesser degree than in QF. In both cases, the presence of lexical/structural effects and lack of semantic ones severely restrict its use. As such, it does not occur freely across all eligible contexts. In the next chapter, one of the few cross-variety differences identified here will be investigated in more detail: the proportional differences in the use of *falloir*.

Chapter 4 THE EXPRESSION OF NECESSITY

4.1 Introduction

The previous chapter focused on subjunctive use in contemporary HF and compared it to what was previously found by Poplack, Lealess & Dion (2013) for QF. The underlying grammar of the subjunctive is strikingly similar across the two varieties despite what appears to be differences in overall rate of occurrence, which was found to result from disproportionate distributions of certain subjunctive governors – notably *falloir*. Further analysis revealed that the proportion of *falloir* was *much* lower in HF than in QF. This drastic proportional difference raises an important question: why is the overall proportion of *falloir* so much lower in HF than it is in QF? This expression represents one of the most common ways of expressing necessity or obligation in French (Lealess 2005; Thibault 1991b); therefore, if the frequency of *falloir* is so low in HF, how exactly are HF speakers expressing this meaning? This question is the driving force behind what is presented in this chapter.

Modality is "the linguistic phenomenon whereby grammar allows one to say things about, or on the basis of, situations which need not be real" (Portner 2009:1), and can encode an array of semantic nuances, including ability, possibility, obligation, and necessity (Depraetere & Reed:269). It can be expressed both grammatically (i.e., through verbal mood) and lexically (e.g., with modal verbs, auxiliaries); however, in French, it "is primarily expressed lexically" (Thibault 1991b:191). While some may argue that alternations between the subjunctive and the indicative constitute a means of expressing modality, the results of the subjunctive analysis

presented in the previous chapter (and those discussed by Poplack and her colleagues) have shown that such alternations are lexically and structurally, not semantically, motivated.

A fairly traditional distinction used in the classification of modality types is that of epistemic and deontic modality (Portner 2009). Epistemic modality revolves around knowledge (Portner 2009), and "concerns what is possible or necessary given what is known and what the available evidence is" (von Fintel 2006:2). Deontic modality, on the other hand, expresses what is "possible, necessary, permissible, or obligatory" (von Fintel 2006:2) based on a set of rules, which can be imposed by a person, moral obligation, or society (Depraetere & Reed:274; Portner 2009).

Interestingly, in many languages, the same lexical items (usually modal verbs) can be used to express both epistemic and deontic modality. Unlike English, which has several modal auxiliaries, French can be argued to possibly have two: *pouvoir 'may'* and *devoir* 'must' (Thibault 1991b), but only the latter is of interest here. *Devoir* can be used both epistemically and deontically, as is shown in (1) and (2). In the first example, *devoir* expresses the fact that it is necessary for the teacher to know how to run their class, while in the second example, the use of *devoir* expresses the speaker's assumption about the age of the referent.

(1) « Un proviseur tient son lycée hein de même qu'un prof ***doit*** savoir tenir sa classe. » (CFPP/MR2/1797)

'A headmaster runs their high school huh the same way a teacher must know how to run their class.'

(2) « …elle ***doit*** avoir trente-cinq ans maintenant… » (CFPP/JP/31)

'She must be thirty-five years old now.'

However, it has been argued that the epistemic/deontic distinction is too simplistic to capture the semantic nuances expressed by different modal expressions (Portner 2009). Other similar approaches (e.g., Brennan 1993) distinguish epistemic modality from non-epistemic modality (often referred to as root modality), but, as Portner (2009:139) points out, root modals do have anything in common with each other "except for not being epistemic", and thus such classifications also fail to capture the complexities of the modality system. Much research has been done in this area in an effort in to identify the semantic subclasses of modality, all while capturing the complexities therein, but thus far, a clear consensus has not been reached[84]. It is not the intention of this study to contribute to this ongoing debate amongst semanticists. Rather, the focus of this analysis will be on one small subcategory of modality: the expression of necessity/obligation that results from a set of rules which are imposed by a specific source or by moral and/or social obligation more generally.

When a French speaker wishes to express necessity or obligation, they can avail themselves of several lexical options, aside from *devoir* (1,2). *Falloir* 'to be necessary/to have to', as shown in (3) and (4), is considered by many (e.g., Lewis 2015; Larreya 2004) to be the most frequent option, along with *devoir*.

(3) « J'ai perdu ma carte 'vélib' donc euh ***faut qu'*** j' m'en refasse. » (CFPP/OK/2325)

'I lost my velib card so um I have to get another one.'

(4) «…vous savez ***faut*** avoir du courage hein des fois. » (ESLO2/OS6/1151)

'You know you have to have courage huh/eh sometimes.'

[84] For example, Portner (2009) proposes three categories: epistemic, priority (subcategorized into deontic, bouletic, and teleological), and dynamic (subcategorized into volitional and quantificational), while Haquard (2009) distinguishes between epistemic, true deontic, and root (including goal-oriented and ability) modality. See Portner for a detailed overview.

However, other options are also available (5-7), including *être obligé* 'to be obliged to', *avoir besoin* 'to need', and *avoir à* 'to have to'[85].

(5) « Quand moi je veux venir chez moi en voiture je ***suis obligé de*** faire des détours infernaux. » (CFPP/GP/478)

'When I want to come home by car I have to take terrible detours.'

(6) « Il y a des gens qui ***ont besoin de*** faire beaucoup d'efforts. » (ESLO1/MD461/562)

'There are some people who have to make a lot of effort.'

(7) « J'***ai*** pas mal de découpage ou de- ou des préparations ***à*** faire[86]. »

(ESLO2/PT45/397)

"I have quite a bit of cutting or of – or some preparations to do."

Recall that, as was outlined in Section 2.1, a linguistic variable refers to different ways of expressing the same thing without resulting in a change in meaning. This study focuses on one subcategory of modality, which is expressed by a subclass of lexical items (*falloir que*, *falloir INF*, *être obligé*, *devoir*, *avoir besoin*, and *avoir à*), that encodes the root meaning 'it is necessary/obligatory that/for' in French. Each of these lexical items can be used in variation with one another when expressing the semantic notion of obligation/necessity[87].

4.2 Background literature

Most of the literature that explores the expression of necessity in French does so from a semantic perspective and usually focuses on *falloir* and/or *devoir* exclusively, often in the

[85] *Être nécessaire que* was excluded at the outset due to its notable infrequency.
[86] Note that there are also instances of *avoir à* where the direct object does not intervene between *avoir* and *à*. The example given in (7) could theoretically also occur as « J'***ai à*** faire pas mal de découpage ou- de ou des préparations ». Both types of structures were included in this study.
[87] From this point forward, these lexical items are referred to as expressions of necessity even though they may also encode obligation.

context of modality or modal verbs in general (e.g., Alba-Salas 2004; Larreya 2004; Baika 1998; Valiquette & Lesage 1995; Portine 1985; Sueur 1981; Bryant 1980; Parret 1976; Benveniste 1974; Huot 1974). Not only are these two expressions of necessity considered to be the most frequent, but they are also said to be stylistically conditioned (e.g., Lewis 2015). It has been suggested that *falloir* is more commonly associated with colloquial or informal speech, while *devoir* tends to be reserved for more formal situations (e.g., Lewis 2015; Labbé & Labbé 2010). In addition, *devoir* is considered to be more insistent (i.e., expresses a stronger degree of obligation) than *falloir* (Larreya 2004). Unlike with the subjunctive, prescriptive grammarians generally do not explicitly favour one variant over another or correlate specific variants with social characteristics, according to an ongoing project at the Sociolinguistics Laboratory at the University of Ottawa. Instead, they tend to mostly focus on *devoir* and *falloir* as the means of expressing necessity in French, without identifying a clear distinction between them.

Even though much work has been undertaken in order to characterize the semantic and syntactic characteristics of *devoir* and *falloir* in HF, not a single study has focused on the necessity system as a whole, including the other possible variants. Moreover, few studies that examine any part of the necessity system do so from a quantitative or variationist perspective. Exceptions include Thibault (1991b), Lealess (2005), Lewis (2015), and Labbé & Labbé (2013), which are briefly outlined below.

Thibault (1991b)

Thibault (1991b) examined a variety of modal expressions, both epistemic and deontic, in her diachronic study of French spoken in Montreal, but dedicated a large portion of her study to the semantic and variationist analysis of *devoir* and *falloir que*. She also identified *être obligé* as

being in variation with other two in most contexts but excluded it from most of the analysis due to "its questionable status […] as a modal operator" (ibid., 210) in certain contexts[88].

She analyzed the internal conditioning of both *falloir* and *devoir,* and found that each variant was conditioned differently: *devoir* was favoured with perfect and present tenses, inanimate, and definite subjects (specifically 1st and 2nd person), while *falloir que* was favoured with future and imperfect tenses and indefinite or impersonal subjects. She also found that *devoir* was favoured in topicalized, relative, prepositional, and interrogative clauses, while *falloir que* was preferred in both complement and independent clauses. Her findings suggest that each expression of necessity is favoured in specific grammatical contexts with very little overlap. She also noted that both male speakers and those of the highest socioeconomic class were more likely to use *devoir* over *falloir*, the latter finding suggesting a possible stylistic difference.

Lealess (2005)

Lealess' (2005) variationist study of the expression of necessity in 20th century QF is one of the most comprehensive to date. In her study, she analyzed speakers from Hull, Quebec and then compared her results to native French speakers from Montreal (adapted from Thibault 1991b) as well as to L2 French speakers from both Montreal and Toronto. Like Thibault, she considered both internal and external conditioning factors and their relative importance in the choice of one expression of necessity over another. Unlike most other researchers, she included not only *falloir* and *devoir*, but also *être obligé*, *avoir à*, and *avoir besoin* and noted their contribution (in terms of rates) to the overall system of the expression of necessity in QF.

[88] Thibault suggests that there are equivalency issues with *être obligé*, compared with *falloir* and *devoir* because the latter two variants both exhibit an important characteristic of modal auxiliaries (i.e., they cannot be negated – only the complement of necessity can be), while *être obligé* does not (i.e, it can be negated). These three variants are included in the present study, but negated contexts are excluded in order to minimize this equivalency issue (see Section 4.3.2 for more).

According to her results, the system in QF is overwhelmingly headed by *falloir*, as is shown in Table 4.1 (adapted from Lealess 2005). The other variants, with the exception of *être obligé* are practically absent from the system.

Variant	**QF (OH)** **% (n)**
Falloir que	58% (429)
Falloir INF	26% (192)
Être obligé	12% (89)
Avoir à	1% (1)
Devoir	1% (9)
Avoir besoin	1% (7)
TOTAL	99% (736)

Table 4.1: Overall rates of expression of necessity variants in QF (adapted from Lealess 2005: 25).

In terms of the internal conditioning of each variant in the Hull sample, Lealess limited her analysis to *falloir que*, *falloir INF*, and *être obligé*. She found that the imperfect tense favoured *falloir INF*, while the past and future favoured *être obligé*. No tense effect was reported for *falloir que*. She also noted that definite subjects favoured *être obligé* while, like Thibault (1991b), indefinite subjects favoured *falloir que*. She found no effect for animacy of the subject but did note that internal sources of obligation favoured *falloir que*, while external ones tended to favour *être obligé*. Since a variationist analysis involves looking at the proportion of each factor associated with each variant, her results are not easily comparable with Thibault's because they did not examine a) the same internal conditioning factors, or b) the same variants (i.e., Thibault: *devoir, falloir que;* Lealess: *falloir que, falloir INF, être obligé*).

Labbé & Labbé (2010)

Labbé & Labbé's (2010) study focused on the use of modal constructions (including those expressing both epistemic and deontic meanings) in HF political speeches and some spoken language. Their sample consisted of 1774 political speeches, from various presidents, totalling approximately six million words. In a list of the twenty most frequent verbs in a subsample of Jacques Chirac's speeches, *devoir* and *falloir* both appeared as modal auxiliaries (i.e., followed by an infinitive verb), although *devoir* was slightly more frequent than *falloir*. In the spoken language sample, they found the opposite: *falloir* was more frequent than *devoir* thus suggesting a stylistic effect. Unsurprisingly, *avoir* and *être* were also included amongst the most frequent auxiliaries in most of the lists; however, there is no way of knowing how many of those occurrences were contained within the other necessity variants of interest to us, such as *avoir besoin, être obligé* and *avoir à.* Moreover, the authors did not stipulate which of the occurrences of *falloir* and *devoir* were expressing a deontic meaning (as opposed to an epistemic one) as these were not differentiated. As such, this study, while quantitative, does not contribute greatly to a clear understanding of the expression of necessity in French. It does, however, reveal that both *devoir* and *falloir* are relatively frequent in both political and spontaneous speech and that they appear to be subject to stylistic conditioning.

Lewis (2015)

Lewis' (2015) quantitative study compared the use of the two most common variants in French (*falloir, devoir*) and English (*have to, must*) in the speech of politicians. She found that *devoir* (71%) and its English equivalent *must* (52%) were actually much more frequent than *falloir* (29%) and its English equivalent *have to* (16%) in the politicians' speeches, which is

contrary to what is claimed to be the case in everyday speech (e.g., Tagliamonte & Smith 2006 for English; Lealess 2005 and Thibault 1991b for French) thus suggesting that *devoir* is more likely to be used in formal contexts, as do Labbé & Labbé (2010).

She also considered some internal factors, including the type of subject, voice, polarity, and negative effect (i.e., "whether the realization of the proposition affects anyone negatively") (Lewis 2015:161) and through mostly qualitative analysis, found that in general, *devoir* was more closely associated with subjective obligation (i.e., "speaker-oriented desirability") and *falloir* with objective obligation (ibid., 170-171). Her result for *falloir* is in opposition to Lealess' (2005); however, it is not clear if she separated *falloir INF* from *falloir que* in her analysis.

Summary

Since most of the studies of the expression of necessity in French are semantic and/or descriptive in nature, rarely consider any expression of necessity other than *falloir* and *devoir*, and hardly ever go beyond overall rate of occurrence, one cannot easily validate claims of frequency, formality, or internal conditioning. Nonetheless, some interesting findings have emerged, including the association of *devoir* with formality and upper classes (e.g., Lewis 2015; Thibault 1991b), as well as the overall frequency of *falloir* compared to the other variants in spoken language (e.g., Lewis 2015; Labbé and Labbé 2010). In addition, despite the fact that the expressions of necessity tend to vary across studies, those who examined internal conditioning (e.g., Lealess 2005; Thibault 1991b) have found that the variability is not at all random, but rather conditioned by specific linguistic and social factors. There is much remaining to be learned about the role of each of the variants in spontaneous HF, as well as if and to what degree their system of expressing necessity differs from that of QF. The aim of the present study is to

fill this gap, in the hopes of revealing why the proportion of *falloir que* is so much lower in HF than in QF, as shown in the previous chapter.

4.2.1 This study

A comparative variationist analysis of the expression of necessity in HF and QF was undertaken in order to determine if there are differences between the two varieties and if so, to what extent these (dis)similarities are present in the underlying conditioning. While the main reason for considering the expression of necessity lies in the proportional differences in the use of *falloir que* in the two varieties, it also represents an area of the grammar that a) has remained largely unexplored in HF, and b) has never been subject to a comparative variationist analysis. This analysis will not only tease apart a cross-variety difference in the subjunctive system (identified in Chapter 3), but it will also shed light more generally on the (dis)similarities that may exist between the two varieties through the investigation of these two separate, but related, morphosyntactic linguistic variables.

Questions similar to those addressed in the analysis of the subjunctive in Chapter 3 will be investigated here:

1) Do HF and QF make use of the same system for the expression of necessity?

2) If so, are there parallels in the linguistic conditioning of the variability in both varieties across time?

If parallels are found between the two varieties, this could suggest that the system used to express necessity is either a shared retention from the source variety (i.e., pre-colonial HF) or a shared innovation that occurred in both. If there are differences, this would indicate that only one

variety underwent linguistic change at some point, which may be revealed by the diachronic analysis.

4.3 Methodology

Lealess' (2005) study of necessity in QF will be used as a 20th century point of comparison for this analysis due to the fact that she has already analyzed this variable from a variationist perspective in the 20th century QF corpus with which 20th century HF has been compared thus far. However, there is no study of the expression of necessity in 21st century QF. In order to fill this gap, a small subsample of 21st century QF (twenty students, aged 15-16) was extracted from *Le français en contexte: milieux scolaire et social* (Poplack & Bourdages 2008), which consists of spoken, informal French[89], and will be subjected to analysis here.

4.3.1 Circumscribing the variable context

Unlike Thibault (1991b), *falloir INF*, *avoir à*, and *avoir besoin* are included here as variants in addition to *devoir, être obligé*, and *falloir que*. Thibault made no mention of *avoir à* or *avoir besoin* in her study, perhaps due to their relative infrequency, and instead focused solely on *falloir que, être obligé, and devoir*, although, as mentioned above, she excluded *être obligé* from the analysis of internal conditioning on the basis that it only expresses the modal value of obligation and not necessity[90]. Thibault (1991b:210) also makes a case for excluding *falloir INF* from the variable context on the basis that it is not equivalent to the other variants. She explains

[89] A first pass at extraction was done as part of an ongoing project at the Sociolinguistics Laboratory at the University of Ottawa. I continued the process and coded all of the data for the subsample used here.

[90] Thibault distinguishes between obligation and necessity in her study, and suggests that while *falloir*, *devoir*, and *être obligé* can be used to express obligation, *être obligé* is unable to express necessity; only obligation. She suggests (based on intuitions) that *être obligé* cannot replace *falloir* and *devoir* in cases where necessity (as opposed to obligation) is expressed. In order to maintain complete objectivity, all lexical items that encode the meaning "it is necessary/obligatory that/for" are included, and as such, no contexts are excluded on the basis of intuitive evidence since, in theory, any of the variants included in this study can be used interchangeably to encode this meaning.

that due to the fact the subject of *falloir* is impersonal (e.g., ***il*** faut…), it cannot be equated with *devoir* unless it is followed by a tensed clause. Examples (8) and (9) are considered equivalent by her, while (10) is not interchangeable with either of them since it has no overt subject that is equivalent to *tu* (examples from Thibault 1991b:210).

(8) Tu ***dois*** faire X. 'You have to do X.'

(9) Il ***faut que*** tu fasses X. 'You have to do X.'

(10) Il ***faut*** faire X. 'X has to be done'

Lealess (2005), on the other hand, did include *falloir INF* as a variant (along with the others included here) in her study of native (Ottawa-Hull) and L2 (Montreal) French, explaining that, while there may be clear arguments for its non-equivalence with the other variants, it is possible that L2 speakers would chose this way of expressing necessity in order to avoid the more complicated option (*falloir que*), which requires a tensed subordinate clause. Using *falloir que* then leads to the choice of what inflectional morphology to apply to the verb, which can be a daunting task for a second language speaker.

In the present study, the variant *falloir INF* is included in the variable context, but not for the same reasons as Lealess since the HF corpora used here consist only of native HF speakers. First, including it enhances the degree of comparability between Lealess' study of QF and this one. Secondly, and more importantly, it can be argued that *falloir INF* does, in fact, belong in the variable context for several reasons. First, while it is true that these constructions have no overt subject, contextual information is often enough to identify the referent, as in (11).

(11) « …Tous les dimanches matin mon père il était très maniaque il avait tout rangé et tout et il prenait son Guide Bleu et il ***fallait faire*** tous les musées les uns après les autres. » (CFPP/MO/3222)

'Every Sunday morning my father was very excited, he put everything away and he took his Blue Guide and we had to do all the museums, one after the other.'

In (11), the speaker is referring specifically to her and her family who had to visit museums every weekend in order to please their father. The same idea could easily be expressed by any of the other variants without triggering a change in the meaning 'we had to do all the museums' (i.e., 'we had to visit all the museums'), as shown in examples (12-16).

(12) « … et il ***fallait qu'***on (def.) fasse tous les musées les uns après les autres. »

(13) « … et on (def.) ***devait*** faire tous les musées les uns après les autres. »

(14) « …et on (def.) ***avait besoin*** de faire tous les musées les uns après les autres. »

(15) « … et on (def.) ***était obligé*** de faire tous les musées les uns après les autres. »

(16) « …et on (def.) ***avait à*** faire tous les musées les uns après les autres. »

It is clear then that despite the absence of an overt subject, *falloir INF's* referent can often be identified using contextual information and this variant can be replaced by any of the other variants without compromising the intended meaning (i.e., expressing some degree of necessity/obligation).

The second reason for including *falloir INF* in the analysis (along with all of the other variants) is that since there are multiple ways in which to express necessity in French, it is important to consider all of them in order to adhere to the Principle of Accountability and crucially, to get a clear understanding of how the entire system necessity works as a whole.

Finally, the disproportionate number of *falloir que* in HF compared with QF in the analysis of the subjunctive was presented in Chapter 3, where the possibility was noted that it may be being used with an infinitival complement instead of a tensed one (labelled by some as a subjunctive replacement strategy). In order to determine whether or not this is the case, it was imperative to include *falloir INF* in the analysis.

Extraction

Every occurrence of each of *falloir que*, *falloir INF*, *être obligé*, *devoir*, *avoir besoin*, and *avoir à* was extracted using a word search function in Concorder Pro. Once each of these forms was extracted, only those followed by an infinitive verb were retained, as opposed to those that were followed by a noun phrase (e.g., *il a besoin d'une chaise* 'he needs a chair') or some other linguistic element. The only exception is *falloir que*, which is followed by a tensed verb in the subordinate clause, which can be indicative, conditional, or subjunctive (as was shown in the previous chapter). Additional exclusions that were necessary in order to adhere to the variable context are described below, as outlined in Thibault (1991b).

4.3.2 Exclusions

All conditional forms of the variants were excluded from the analysis since they tend to encode a meaning other than 'it is necessary/obligatory that/for', or at least do not do so with the same degree of strength as the other modal expressions under analysis here. The equivalent of these forms in English is 'ought' or 'should', which semanticists agree are intuitively different, and perhaps less sensitive to the rules underlying the interpretation of 'strong' necessity modal

expressions (von Fintel & Iatridou 2008)[91]. In examples (17) and (18), it is clear that it is a suggestion that is being expressed rather than an obligation.

(17) « Je crois qu'on ***devrait*** leur apprendre surtout la morale. » (ESLO/103/532)

'I believe we should teach them morality above all.'

(18) « Alors ma petite dame ***faudrait*** peut-être vous dépêcher un peu. » (CFPP/NN/133)

'So my dear maybe you should hurry up a little.'

Some of the variants can be used to express epistemic modality (i.e., "what is possible or necessary given what is known and what the available evidence is") (von Fintel 2006:2). Only instances of the variants which express root necessity or obligation are included, while those which express assessments about probability or possibility, as in (19) and (20) with *devoir*, are excluded.

(19) «…l'espagnol ***doit*** être plus dur au point de vue grammatical.» (ESLO1/RC196/602)

'Spanish must be harder from a grammatical point of view.'

(20) « …elle ***doit*** avoir trente-cinq ans maintenant… » (CFPP/JP/31)

'She must be thirty-five years old now.'

Similarly, the expression *avoir à* does not always express necessity or obligation. There are many cases, like in (21) and (22), where it encodes an epistemic meaning of possibility or probability. These are also excluded from the analysis.

(21) « …bah y ***a*** beaucoup ***à*** faire à Orléans hein. » (ESLO1/FC716/243)

'bah there's lots to do in Orléans eh'

[91] See von Fintel & Iatridou (2008), Copely (2006), and Bybee, Perkins & Pagliuca (1994) for further discussion on the classification and interpretation of these forms.

(22) «…vous savez, y ***a*** beaucoup ***à*** dire là-dessus. » (ESLO1/JO304/906)

'You know, there's a lot to say about it.'

Negated occurrences lead to equivalency problems (Thibault 1991b) and thus, these were not retained for analysis in the present study. For example, with *falloir* (23) and *devoir* (24) it is the complement of necessity that is negated (i.e., the action expressed by the verb that co-occurs with the variant) while with *être obligé, avoir à*, and *avoir besoin* (25), the obligation (i.e., the variant) is negated (Lealess 2005; Thibault 1991b).

(23) « ***Faut*** pas y aller avec des talons. » (ESLO2/OS6/1089)

'You can't go there with heels.'

(24) « Maman nous disait toujours on ne ***doit*** pas mettre ça quand on écrit. » (ESLO1/JO304/1257)

'Mom always told us you can't put that when you write.'

(25) « …il trouvait même que ***j'avais*** pas ***besoin*** d'attacher de l'importance à tout ça lui. » (ESLO1/DT78/1363)

'He even found that I didn't have to attach importance to all that.'

Finally, all occurrences of the variants that filled a discursive function (e.g., *il faut dire que, je dois dire* 'I must say…', etc.) as well as any otherwise ambiguous occurrences which did not clearly express "it is necessary/obligatory that/for" were excluded from the analysis. After the extraction and exclusion process, a total of 2046 tokens in 20th and 21st century HF, and 510 in 21st century QF (along with Lealess' 793 tokens for 20th century QF) were retained for analysis.

It should be noted that the tokens of *falloir que* examined here include most of those analyzed in the previous chapter, along with many more. Recall that many of the tokens had to be excluded from the subjunctive variable context on the grounds that the embedded verb form was ambiguous (i.e., identical indicative and subjunctive forms). These tokens are included here due because the ambiguity of the tense of the embedded verb has no effect on its use as an expression of necessity. In addition, a small number of the tokens excluded here were included in the subjunctive analysis (Table 4.2). As a result, more *falloir que* tokens are analyzed here and these include most of those considered in Chapter 3.

	Subjunctive (Chapter 3)	**Necessity (Chapter 4)**
Ambiguous embedded verb	X	√
Unambiguous embedded verb	√	√
Conditional matrix	√	X
Negated matrix	√	X
Total tokens included	**348**	**478**

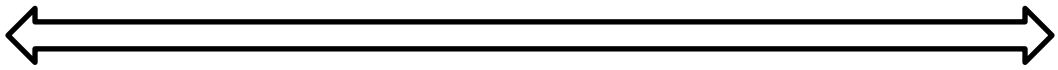
Table 4.2: Comparison of number of *falloir que* tokens considered in the variable context of the subjunctive and the expression of necessity in 20th and 21st century HF.

4.3.3 Factors under analysis

The factor groups tested in this analysis are, for the most part, based on Coates' (1983) cline of strength of obligation (encompassing root necessity/obligation)[92]. In a detailed discussion of the modals of obligation and necessity embedded in a larger discussion of the

[92] This model of strength of obligation is one of many. More recent models are set within Kratzer's framework (1991) where models "quantify over a set of worlds that is calculated from a modal base of accessible worlds and an ordering source which ranks the worlds in the modal base" (cited in von Fintel & Iatridou 2008:117). For example, Kratzer suggests that "*p* is a weak necessity iff *p* is a better possibility than *not p*", while Copely (2006) proposes that a prejacent must "be true in *all* favoured worlds" in order for a modal to encode strong necessity (cited in von Fintel & Iatridou 2008:118). Coates' (1983) model of strength of obligation informs the present analysis because it was most easily operationalizable, and because it has inspired other sociolinguistic studies of deontic modality (e.g., Tagliamonte & D'Arcy 2007; Tagliamonte & Smith 2006). It should be noted, however, that the goal of the present study is not to identify what constitutes strong/weak necessity in French, but rather to situate the variants in relation to one another according to specific linguistic factors, and most importantly, compare their conditioning across two varieties of French. Future studies that marry these recent semantic approaches to modal strength with quantitative analysis would contribute greatly to this area of research.

entire English modal auxiliary system, Coates (1983:32) proposed a cline, ranging from strong (encoding a meaning of 'it is imperative that') to weak (encoding a meaning closer to 'it is important that') obligation. In other words, not all expressions of necessity encode the same strength of obligation, but crucially, they all encode the basic meaning of "it is necessary for" (Coates 1983:32). A prototypical expression of strong obligation includes four factors, which are shown in Figure 4.1.



[-STRONG]	**[+STRONG]**
- Inanimate subject	- Animate subject
- Stative main verb	- Non-stative (activity) main verb
- Low level of speaker involvement	- High level of speaker involvement
- Subject not under authority of speaker	- Subject under authority of speaker

Figure 4.1: The features included in a prototypical expression of strong obligation (cited directly and in some cases adapted from Coates 1983:33).

Coates (1983) also correlated the strength of the obligation to the grammatical person of the subject, with utterances containing second person subjects usually exhibiting a stronger degree of obligation than first person subjects, followed by third person ones, based partially on the level of speaker involvement. That is, the strength of the demand is usually stronger if the speaker is obliging the interlocutor to do something than if the speaker is reporting on a third party. As Tagliamonte & Smith (2006:359) point out, this only applies to definite subjects since indefinite ones are incapable of "encoding a strong reading of 'it is imperative that'".

Most expressions of necessity are not prototypical expressions of strong or weak obligation; therefore, they fall at various places along this proposed continuum. As Coates

(1983) points out, an utterance cannot be easily placed along the continuum unless it falls under one of the two extremes since the area in between is considerably 'fuzzy' (ibid., 34). Despite the fact that some expressions of necessity may encode a stronger sense of obligation than others, the important point to make here is that they always express some degree of necessity/obligation and thus constitute variants of a linguistic variable. While Coates' (1983) cline is based on modal auxiliaries in English, the same [+strong] diagnostics she proposed can also apply to French, which is what is done in the present study with some of them (those that could be tested objectively). Each of the factors groups considered in this study is outlined below.

4.3.3.1 External (social) factors

The same three social factors that were examined in relation to subjunctive use in Section 3.3.3.1 were also included in the analysis of the expression of necessity. Once again, their inclusion in the analysis is largely exploratory. They include sex, age, and point in time.

4.3.3.2 Internal (linguistic) factors

In addition to the external factors listed above, each token was also coded for several internal (linguistic) factors in order to determine which, if any, condition the expression of necessity in HF.

Subject of the complement of necessity

The subject of the complement of necessity refers to the grammatical subject of the expression of necessity[93]. The subject of the complement of necessity was coded according to both grammatical person and definiteness. 1st person subjects are always definite, but 2nd person

[93] For example, each of the following would be coded as first person: ***je*** *suis obligé de le faire,* ***j'****ai x à faire,* ***j****'ai besoin de faire x,* ***je*** *dois faire x, il faut que* ***je*** *fasse/fais x.*

subjects (*tu/vous*) can be either definite or indefinite. 3rd person subjects, which include pronouns and nouns, can also be definite or indefinite and were coded accordingly. Certain subject types required careful review of the context, particularly 2nd person (*tu/vous*) subjects and *on*, many of which are actually ambiguous when it came to definiteness. Examples of each are given below (Table 4.3). The coding of the subject of the complement of necessity applied to all variants except for *falloir INF*, which only has an expletive subject.

In terms of previous findings, Thibault (1991b) found that definite subjects favoured *devoir,* while indefinite subjects favoured *falloir que*. Lealess found the same result for *falloir que* in her study of QF and also noted that *être obligé* was more likely to co-occur with a definite subject.

Subject type	**Example**
1st person definite (*je*)	« J***'***ai ***besoin d***'aller boire un thé à la menthe. » (ESLO2/BC41/300) 'I have to go drink a mint tea.'
2nd person definite (*tu/vous*)	« J'ai dit ma fille ***il faut*** absolument ***que*** que tu sortes de ta coquille. » (ESLO1/JI306/1075) 'I said my daughter, you absolutely have to to come out of your shell.'
2nd person indefinite (*tu/vous*)	« Si tu travailles comme moi t'***es obligée*** d'y aller le week-end.» (CFPP/BD/3503) 'If you work like me, you have to go there on the weekend.'
3rd person definite (*il/elle/N/on/ils/elles*)	« Parce que l'aînée a été très très malade elle a été ***obligée de*** lâcher son travail. » (ESLO1/ZD585/59) 'Because the oldest was very sick she had to quit her job.'
3rd person indefinite (*il/elle/N/on/ils/elles*)	« ***Faut que*** tout le monde travaille. » (ESLO1/FC716/134) 'Everyone has to work.'

Table 4.3: Examples of the different types of subject of the complement of necessity.

This factor group was eventually separated into two groups for the purpose of the analysis. First, all of the data were collapsed into two categories (definite vs. indefinite), as a

separate factor group, with the former being more likely to encode a stronger sense of obligation than the latter (Coates 1983). Second, in contexts where the expression of necessity was more likely to occur with a definite subject or referent, grammatical person (i.e., 1st, 2nd, and 3rd person definite) was coded since the strength of the obligation can be expressed this way according to Coates. The proposed strength of each is illustrated in Figure 4.2.

Figure 4.2: Relative strength of each type of grammatical person in the expression of necessity (adapted from Coates 1983:37).

Since the variant *falloir INF* has no overt subject, it was not possible to code it in the same manner as the others for this factor group. However, it was important to find a way to capture the perceived generic nature of this variant. Therefore, the referential subject of each token of *falloir INF* was coded as 'definitely non-generic' or 'potentially generic'. Examples are given in Table 4.4.

Subject type	**Example**
Definitely non-generic (definite)	« …tous les dimanches matin mon père il était très maniaque il avait tout rangé et tout et il prenait son Guide bleu et il ***fallait faire*** tous les musées les uns après les autres. » (CFPP/MO/3222) 'Every Sunday morning my father was very excited, he put everything away and he took his blue Guide and we had to do all the museums, one after the other.'
Potentially generic/indefinite	« J'estime il ***faut être*** quand même assez franc. » (ESLO1/GS117/281) 'I believe we have to be quite frank anyway.'

Table 4.4: Examples of the generic vs. non-generic nature of the null subjects in *falloir* INF tokens.

Tokens coded as 'definitely non-generic' were those that could only have a non-generic reading (i.e., they have a definite referential subject even though it is not expressed overtly and cannot have a generic reading). In the example in Table 4.4, the speaker is referring to the fact that her family had to visit the museums every Sunday. There is no possible generic reading available here.

The tokens that were coded as 'potentially generic' were those that could potentially have a generic reading. This includes those that could only be generic and also those that could have a generic or specific reading (i.e., are ambiguous). These tokens, along with the other variants, were later recoded as 'definite' (definitely non-generic) and 'indefinite/generic' (potentially generic) in order to ensure comparability.

Hypothesis: Strong necessity is expressed in part by the use of definite subjects over indefinite ones and within definite subjects, by 2nd and 1st person subjects over 3rd person subjects. Since 'devoir' is hypothesized to be the 'stronger' variant compared with 'falloir' (Larreya 2004), it can be expected that definite subjects will favour 'devoir', particularly 2nd and 1st person ones, and indefinite subjects will favour 'falloir'. Definite subjects will also favour 'être obligé' if the findings for HF are in line with Lealess' (2005).

Animacy of the subject of the complement of necessity

The subject of the complement of necessity was coded as either animate or inanimate (Table 4.5). This factor group was not coded for the variant *falloir INF* since there is no overt subject[94].

Dubois (1969:119, cited in Huot 1974) claims that in QF, *devoir* with the semantic reading of obligation is associated with animate subjects, while *devoir* with a probability reading

[94] In addition, all referential subjects would necessarily be animate.

is more likely to be associated with inanimate subjects. Thibault's (1991b) results indicated otherwise, with inanimate subjects favouring *devoir*. Lealess (2005) found no effect on variant choice for this factor group.

Subject type	**Example**
Animate	« Un proviseur tient son lycée hein de même qu'un prof ***doit*** savoir tenir sa classe. » (CFPP/MR2/1797) 'A headmaster runs their high school huh the same way a teacher has to know how to run their class.'
Inanimate	« Je trouve encore une fois que la priorité ***doit*** être l'éducation. » (ESLO2/ZD13/829) 'I think once again that education has to be the priority.'

Table 4.5: Examples of the animate and inanimate subjects of the complement of necessity.

Hypothesis: Expressions of necessity with animate subjects fall at the [+strong] end of Coates' (1983) cline, and since 'devoir' is considered to express a stronger sense of obligation than 'falloir', it follows that 'devoir' should be favoured by animate subjects. However, conflicting previous results suggest that this factor group may not be revelatory.

Type of verb

The type of verb of the complement of necessity was coded as either a non-stative (i.e., action) verb or a stative verb, with the former constituting a [+strong] factor according to Coates' cline. Examples are given in Table 4.6. This factor group applies to all of the variants and has not been previously tested in any study of necessity in any variety of French[95].

[95] To my knowledge.

Verb class	Example
Stative	« Le gars quand même euh ***doit*** savoir ce qu'il a à faire. » (ESLO1/PW85/1034) 'Anyway the guy has to know what he has to do.'
Action	« Je lui demande comment je ***dois*** écrire pour les lettres de premier de l'an... » (ESLO1/BX11/554) 'I ask him how I have to write for the New Year's letters.'

Table 4.6: Examples of different types of verb of the complement of necessity.

Hypothesis: If the strength of obligation factors proposed by Coates are supported, action verbs should favour 'devoir' and stative verbs should favour 'falloir' based on Larreya's (2004) claim that 'devoir' is more insistent than 'falloir'.

Source of the obligation

Following Thibault (1991b) and Lealess (2005), each token was coded according to the source of the obligation. An internal (unrestricted) source of obligation refers to the fact that the obligation is self-imposed (e.g., in cases of moral obligation), whereas an external (restricted) source of obligation originates from a source external to the person on whom the obligation is being imposed (e.g., rules or regulations) (Thibault 1991b:208)[96]. As was the case in the studies of Thibault (1991b) and Lealess (2005), many of the tokens coded in this study (approximately one third) were ambiguous with regards to source of obligation and were thus excluded from this portion of the analysis. Examples of each are provided in Table 4.7.

Sueur (1981) claims that when the source of obligation comes from an internal source, *devoir* can be replaced by *falloir que* and when the source of obligation comes from an external source, it should be replaced by *être obligé*. However, both Thibault (1991b) and Lealess (2005)

[96] This differs from the distinction between subjective and objective obligation commonly referred to in the literature (e.g., Huddleston & Pullum 2002), which focuses on the level of speaker involvement. While subjective sources tend to be associated with stronger degrees of obligation (Coates 1983), there is no such hypothesis for the internal/external distinction.

found cases of internal obligation with *être obligé* and cases of external obligation with *falloir que*. Lealess confirmed that her results patterned in the manner suggested by Sueur (1981) in that internal obligation favoured *falloir que*, but did not find much of an effect with the other variants.

Verb class	**Example**
External	« Mon mari m'a ***obligée à*** travailler.» (ESLO1/IC346/154) 'My husband made me work.'
Internal	« On a une langue on ***doit*** conserver.» (ESLO1/UN412/630) 'We have a language that we have to preserve/conserve.'
	« Je pense que quand on l'écrit on ***doit*** faire attention à à l' orthographe. » (ESLO1/NS571/568) 'I think that when we write, we have to pay attention to- to spelling.'

Table 4.7: Examples of tokens with internal and external sources of obligation.

Hypothesis: If QF and HF exhibit underlying similarities, the results for HF should resemble Lealess' in that internal sources of obligation favour 'falloir que'.

Each of these factor groups will be tested in order to determine their relative contribution to the expression of necessity in HF. If a variant is associated with weak obligation, it would be expected to correlate with the [-strong] factors, situated at the weak end of the cline (i.e., indefinite, inanimate subjects, and stative verbs). Conversely, if a variant encodes strong obligation, a correlation with [+strong] factors (i.e., definite (2^{nd}, 1^{st}), animate subjects, and non-stative verbs) would be likely.

It should be noted that most of the hypotheses focus on claims about *devoir'*s 'insistence' and previous findings from Lealess (2005) and Thibault (1991b). Since this study is the first to analyze the entire system of the expression of necessity in HF and the role of the strength of obligation in variant choice, the analysis of these factor groups (particularly when it comes to

those other than *falloir* and *devoir*) is largely exploratory and will hopefully shed light on the role that each plays in the system as a whole. The results of the analysis are presented in the following section.

4.4 Results

In the ensuing sections, the expression of necessity in 21st century HF is compared to 21st century QF in an attempt to identify any cross-variety (dis)similarities. A diachronic analysis follows, in order to determine if any changes took place between the late 20th century and the early 21st century in HF. Finally, the results are compared to the diachronic trajectory of the expression of necessity in QF, which will be established by comparing the results of 21st century QF presented here to Lealess' (2005) analysis of 20th century QF (where possible).

4.4.1 Synchronic comparative analysis of the expression of necessity in HF and QF[97]

4.4.1.1 Overall rates

The results presented in Table 4.8 show the overall rate of the different necessity variants in contemporary HF and confirm one fact that has been repeatedly alluded to in the literature (e.g., Thibault 1991b): *falloir* is one of the most common variants. However, these findings do not support the claim that *devoir* is the other most frequent variant. In fact, *falloir* (including *falloir INF* and *falloir que*) is used 73% of the time, which is far more than any other variant. *Devoir*, *être obligé*, and *avoir à* are all used with approximately the same frequency: 7-10% of the time. *Avoir besoin*, however, is rarely used at all (3%).

[97] Here, the focus is on 21st century HF instead of 20th century (as was the case for the subjunctive) because I continued the extraction and proceeded with the analysis for 21st century QF myself. All results for 20th century QF are adapted from Lealess (2005). No reanalysis of her results was conducted, with the exception of the re-extraction of *avoir à*, which affects only the overall rate and distribution of the variants, albeit minimally.

Variant	N	%
falloir INF	389	49
falloir que	193	24
devoir	77	10
être obligé	59	7
avoir à	63	8
avoir besoin	21	3
TOTAL	**802**	**101**[98]

Table 4.8: Overall rate of use of each variant in the expression of necessity in 21st century HF[99].

Taking a closer look at *falloir*, the results show that *falloir INF* is used twice as often as *falloir que* and that *falloir que* only represents about one quarter of the overall data pool. These two variants are rarely considered separately from one another in the literature. Recall that *falloir INF* is followed by an infinitive verb (26), while *falloir que* is followed by a tensed embedded clause (27), which almost always selects an embedded verb with subjunctive morphology, as was shown in Chapter 3.

(26) « Je pense qu'il ***faut*** faire attention au contenu des mots. » (ESLO2/AU49/2058)

'I think you have to pay attention to the content of the words.'

(27) « Je pense qu'il ***faut que*** ça évolue. » (ESLO2/AU49/2053)

'I think that it has to evolve.'

How do these findings compare to variant distribution in 21st century QF? There are two clear similarities to take note of (Table 4.9). First, *falloir* is the majority variant in both varieties but accounts for a larger proportion (88%) of the data in QF than in HF (73%). Second, both varieties make use of the same inventory of variants to express necessity and all variants other than *falloir* play a minimal role, each representing between 1% and 10% of the system.

[98] Due to rounding.

[99] As was the case in the subjunctive analysis in Chapter 3, both sets of 21st century HF data are combined under one label "21st century HF".

	HF		QF[100]	
	N	**%**	**N**	**%**
falloir INF	389	49	42	8
falloir que	193	24	406	80
devoir	77	10	15	3
être obligé	59	7	26	5
avoir à	63	8	18	4
avoir besoin	21	3	3	1
TOTAL	**802**	**101**[101]	**510**	**101**[102]

Table 4.9: Overall rate of use of each variant in the expression of necessity in 21st century HF and QF.

Thus far, the main difference between the two varieties lies in their differential use of *falloir* for expressing necessity. In QF, it is *falloir que* that dominates (80%), while *falloir INF* is relatively rare (8%) and is used about as often as the other variants. In HF, on the other hand, it is *falloir INF* which accounts for the lion's share of the data (49%), while *falloir que* is used much less frequently (24%), but still at least more than twice as often as any of the other variants.

These findings may help to explain the disproportionate number of *falloir que* tokens in subjunctive-selecting contexts across the two varieties: when expressing necessity, QF speakers chose *falloir que* (the only variant which co-occurs with a tensed – usually subjunctive-embedded verb) three times more often than HF speakers do[103]. In other words, more than three quarters of the time (76%), HF speakers will use a variant other than *falloir que* to express necessity and in doing so, a highly favourable subjunctive-selecting context is lost.

[100] Only students (young speakers) are included in the 21st century QF data sample. While the corpus does contain older speakers, they are French teachers and thus are unlikely to represent the community norm due to an increased level of formality in their speech.
[101] Due to rounding.
[102] Due to rounding.
[103] Recall that the *falloir que* tokens examined here also include the contexts which were excluded as ambiguous from the analysis of the subjunctive; however, the proportion of tokens excluded on these grounds was approximately the same in both varieties (40-50%).

4.4.1.2 Internal conditioning

In this section, the internal conditioning of each of the variants used to express necessity is examined in HF and then compared to QF. Where data is sufficient, the statistical significance of the contribution of each factor on variant choice is tested. Results for HF are presented in Table 4.10. Factors which are considered to encode strong obligation as per Coates (1983) are annotated with '+'.

The first factor group (definiteness) considers the definite nature of the subject (or referent in the case of *falloir INF*) of the complement of necessity. Indefinite/generic reference favours *falloir INF*, while definite subjects favour all of the other variants, with the exception of *avoir à*, where no effect is found for this factor group. External sources of obligation and non-stative verbs favour *avoir à* and *être obligé*, while internal sources and stative verbs appear to favour *avoir besoin* and *falloir* (*INF/que*). *Devoir* is favoured by external sources of obligation and stative verbs, though the latter is not selected as statistically significant. Finally, with regards to animacy, inanimate subjects favour *avoir à*, while animate ones favour the remaining variants, with the exception of *devoir* where there is no effect.

Strength	Corrected mean N	*Devoir* .080 77/802			*Avoir à* .054 63/802			*Être obligé* .052 59/802			*Avoir besoin*[104] 3% 21/802			*Falloir que* .190 193/802			*Falloir INF*[105] .455 389/802		
		FW	%	N	FW	%	N	FW	%	N	FW	%	N	FW	%	N	FW	%	N
	Definiteness																		
+	Definite	.62	14	46/342	[ ]	8	28/342	.66	12	42/342	/	5	16/342	.80	46	157/342	.17	16	53/342
	Indefinite/generic	.41	6	28/453	[ ]	8	34/453	.38	4	16/453	/	1	4/453	.26	8	35/453	.77	74	336/453
	Range	*21*						*28*						*54*			*60*		
	Animacy[106]																		
+	Animate	/	18	62/341	/	10	35/341	/	16	57/341	/	6	19/341	/	49	168/341	/	n/a	n/a
	Inanimate	/	18	12/66	/	41	27/66	/	2	1/66	/	2	1/66	/	38	25/66	/	n/a	n/a
	Obligation																		
n/a	Internal	.42	6	18/284	.34	3	9/284	.33	4	10/284	/	5	14/284	.55	29	83/284	.59	53	150/284
n/a	External	.59	12	32/258	.67	12	30/258	.68	15	38/258	/	0	1/258	.45	25	64/258	.40	36	93/258
	Range	*17*			*33*			*35*			/			*10*			*19*		
	Type of verb																		
+	Non-stative	[ ]	9	59/643	.57	9	59/643	.56	9	56/643	/	2	12/643	.46	23	150/643	/	48	307/643
	Stative	[ ]	11	18/159	.25	3	4/159	.26	2	3/159	/	6	9/159	.65	27	43/159	/	52	82/159
	Range				*32*			*30*						*19*					

Table 4.10: Internal factors conditioning the use of all variants expressing necessity in 21st century HF (shading indicates favourable factors to variant choice)[107].

[104] These data were not subjected to a multivariate analysis due to the limited number of tokens.

[105] Type of verb was excluded from the multivariate analysis of *falloir INF* due to an interaction with definiteness: there are only two tokens with a definite referential subject and a stative verb.

[106] Only marginals are provided for this factor group as there was a clear interaction with definiteness due to a lack of definite, inanimate tokens.

[107] In this and subsequent analyses, any tokens that could not be objectively coded were excluded from the analysis. Source of obligation is the factor group most affected, with a substantial number of tokens that are ambiguous in this regard.

Since most of these factor groups can be used as diagnostics of the strength of the expression of necessity, it is possible to test whether or not specific variants are more likely to encode a [+strong] meaning. While Coates (1983) suggests that the area between the two ends of the cline tend to be 'fuzzy', the hypothesis put forth here is that the more [+ strong] factors that favour a variant, the stronger the degree of necessity/obligation.

Coates (1983) suggests that when it comes to definite subjects, the grammatical person plays a role in the strength of the obligation, with 2nd person definite subjects more likely to be associated with [+strong] obligation than 1st person ones, while 1st person subjects are associated with stronger degrees of obligation than 3rd person definite ones. The above analysis revealed that definite subjects favour four of the six variants: *falloir que*, *être obligé, devoir,* and *avoir besoin*. But it is unclear if there are differences in the type of definite subject used in each context. This is tested below in the hopes of further distinguishing between them (Table 4.11).

	Falloir que		***Avoir besoin***		***Être obligé***		***Devoir***	
	%	N	%	N	**%**	**N**	%	N
Type of subject								
2nd person def.	71	10/14	7	1/14	0	0/14	7	1/14
1st person	62	90/145	5	7/145	15	22/145	8	12/145
3rd person def.	44	57/130	6	8/130	15	20/130	25	33/130

Table 4.11: Summary of variant choice according to type of definite subject in 21st century HF.

The results presented above reveal that 2nd person definite subjects favour *falloir que*, while 3rd person definite subjects favour *devoir.* With regards to type of subject, one could argue that this places *falloir que* closer to the [+strong] end of the cline, in opposition to *devoir*. However, given the lack of 2nd person definite subjects used with all the variants (n= 14), these findings should be taken with a grain of salt. Since the data were collected using sociolinguistic interviews, it is not surprising to find a lack of such contexts. In fact, it is expected that it would

be rare (and the numbers confirm this) to find the speaker obliging the interviewer to do something. In fact, most of these tokens occur within the context of reported speech, where the speaker is recounting telling a specific person to do something. Both 1st and 3rd person subjects favour *être obligé*, which places it somewhere in between the other two. There is no effect for *avoir à*.

Now that the effects of each factor group on variant choice have been presented, the next step involves assessing what this means in terms of strength of obligation and variant choice more generally. First, with regards to the strength of obligation, Table 4.12 outlines how many [+strong] factors are associated with each variant. It is clear that each one (with the exception of *falloir INF* and *être obligé*) is associated with at least one [+ strong] factor as well as at least one [-strong] one, which reveals little about where each falls on Coates' (1983) cline. It appears that most of the variants can be simultaneously favoured by both strong and weak factors, and, therefore, fall somewhere within the 'fuzzy' part of the cline between the two extremes. Based solely on the number of [+strong] factors favouring them, it would appear that *être obligé* encodes a stronger degree of obligation than the others[108].

[+strong] necessity	*Être obligé*	*Avoir à*	*Devoir*	*Avoir besoin*	*Falloir que*	*Falloir INF*
Animate subject	+	-		+	+	n/a
Definite subject	+		+	+	+	-
Non-stative verb	+	+	-	-	-	-

Table 4.12: Summary of variants according to their degree of obligation: [+] indicates that the factor associated with strong obligation favours the variant, [-] indicates that the factor associated with weak obligation favours the variant, [n/a] indicates that the factor group was not tested.

Despite revealing little about variant use according to the strength of the obligation, this analysis of the internal conditioning of variant choice is still very revelatory in that it outlines a

[108] The analysis of *avoir besoin* is only based on twenty tokens so this result should be taken with a grain of salt.

clear structural context in which each variant is most likely to be selected, which is described below.

Starting with *avoir à*, it should be noted (Table 4.10) that this variant differs from the others due to the fact inanimate subjects of the complement of necessity favour its selection. Further analysis revealed that impersonal subjects tend to co-occur with *avoir à*, as in (28). In fact, nearly two thirds of all the impersonal subjects used in the variable context are found with this variant (61%).

(28) « …y ***a*** le Voyageur Imprudent aussi ***à*** lire par tout le monde. » (ESLO2/AU49/2255)

'There is also the *Voyageur Imprudent* that everyone has to read/to be read by everyone.'

In addition, *avoir à* is most likely to occur with a direct object intervening between *avoir* and *à*, as in (29), as opposed to cases where the object appears before (30) or after (31). Another finding is that *faire* (29,30) plays a notable role compared to other verbs (Table 4.13) in that more than half of the occurrences of *avoir à* contain *faire* as the verbal head of the complement of necessity.

(29) « J'***ai*** encore six ans **à** faire vous voyez. » (ESLO1/DT78/130)

'I still have six years to do you know.'

(30) « Enfin je sais pas les choses que j'***ai à*** faire. (ESLO1/EM229/361)

'Well I don't know the things that I have to do.'

(31) « Il continue lui ses études sans ***avoir à*** supporter euh les choses qui étaient pas bien. » (ESLO2/MF363/644)

'He continues his studies without having to deal with the things that weren't good.'

	%	N
Direct object appears as intervening material	66	38/58
Complement of necessity verb *faire* in these constructions	58	22/38

Table 4.13: Rate of intervening direct object and *faire* in *avoir à* constructions.

In sum, it can be concluded that this variant is largely restricted to the construction shown in (32).

(32) $X_{\text{inanimate}}$ avoir $X_{\text{direct object}}$ à $X_{\textit{faire}}$

Of the remaining variants, *falloir INF* is by far the preferred method of expressing indefinite necessity in 21st century HF, as shown in Table 4.14, which breaks down all tokens expressing indefinite necessity according to variant. The vast majority (74%) of all tokens under any variant expressing indefinite/generic reference fall under *falloir INF*. Recall that since there are no overt subjects used with *falloir INF,* this variant was coded for definiteness by examining each of the tokens within their respective context and determining whether the intended subject could possibly be generic. Tokens that could not have a generic subject (based on contextual evidence) were later recoded as 'definite' (definitely non-generic) and those that could were coded as 'indefinite/generic' (potentially generic). With the other variants, indefinite reference is generally signalled by the use of an indefinite noun or pronoun (*on, tu, vous, ils*) in addition to contextual information.

	N	%
Falloir INF	336	74
Falloir que +indef. (pro)noun	35	8
Other variant +indef. (pro)noun	82	18
Total tokens with indefinite reference	453	100

Table 4.14: Display of the different methods employed in 21st century HF to express indefinite/generic necessity.

With the other variants (*devoir*, *être obligé*, and *falloir que*)[109], a definite subject is selected as a significant conditioning factor in each case, but they are associated with a unique set of underlying conditioning factors which differentiates them from one another. Strength of obligation may also play a role in distinguishing *être obligé* from the others, since *être obligé* appears to encode a stronger degree of obligation.

	Falloir que	***Être obligé***	***Devoir***
Definiteness of subject/referent	Definite	Definite	Definite
Animacy	Animate	Animate	
Source of obligation	Internal	External	External
Type of verb	Stative	Non-stative	Stative

Table 4.15: Summary of internal conditioning factors for *falloir que, devoir,* and *être obligé* in 21st century HF [110] (shading indicates [+strong] factors)**.**

In sum, variant choice for expressing necessity in 21st century HF can be distinguished according to various internal conditioning factors. While these factor groups were tested with the intention of placing each variant along a continuum of strong to weak obligation, little evidence was found to support a clear distinction due to the fact that most of the variants are associated with both [+strong] and [-strong] characteristics and the difference between them on these grounds is minimal. In addition, the analysis of the effects of the factor groups on variant choice

[109] Excluding *avoir besoin* due to lack of data.
[110] *Avoir besoin* is excluded due to lack of data (n= 20) and *avoir à* is also excluded due to its restricted use in a fixed construction.

revealed a structured system of the expression of necessity, with each variant occupying its own niche (summarized in more detail below).

Comparison with QF

Now that the role of each variant in the HF expression of necessity system has been established, these results will be compared with QF in order to see if they share a single underlying grammar for each variant. Results for QF are presented in Table 4.16.

Recall that in terms of variant distribution, the main difference between HF and QF was in the proportion of *falloir: falloir que* represents the majority variant in QF (80%), while *falloir INF* is the majority variant in HF (49%). Due to the lack of data for all other variants in QF, the results of a multivariate analysis are only presented for *falloir que*. The internal conditioning for the remaining variants is provided using marginal results. Similar to HF, where *falloir que* is favoured by definite animate subjects, stative verbs, and internal sources of obligation, in QF, internal sources of obligation and animate subjects favour its use (Table 4.16). There is no effect with the other factor groups (definiteness or verb type) in QF which is unsurprising given that this variant dominates most of the system by accounting for 80% of it.

Falloir INF is favoured in indefinite/generic contexts and by stative verbs in both varieties (Table 4.10, 4.16), despite it being used quite infrequently in 21st century QF (8%) when compared with 21st century HF (49%). *Être obligé* is conditioned nearly identically across the two varieties: definite, animate subjects, external sources of obligation, and non-stative verbs favour its selection[111].

[111] The effect of verb type is not selected as significant in QF, but the marginals point to the same direction of effect, albeit minimally.

Strength		*Avoir à*		*Devoir*		*Être obligé*		*Falloir que*			*Falloir INF*	
	%	4%		3%		5%		80%			8%	
	N	18/510		15/510		26/510		406/510			42/510	
		%	N	**%**	**N**	%	N	FW	%	N	%	N
	Definiteness											
+	Definite	3	7/218	3	6/218	8	18/218	[ ]	80	174/218	5	11/218
	Indefinite/generic	4	11/292	3	9/292	3	8/292	[ ]	80	232/292	11	31/292
	Animacy											
+	Animate	3	14/448	3	15/448	6	25/448	/	87	391/448[112]	n/a	n/a
	Inanimate	21	4/19	0	0/19	0	0/19	/	79	15/19	n/a	n/a
	Obligation											
n/a	Internal	2	3/162	3	4/162	3	5/162	.59	86	139/162	6	10/162
n/a	External	5	12/249	2	6/249	7	17/249	.44	77	191/249	8	21/249
	Range							*15*				
	Type of verb											
+	Non-stative	4	18/462	3	13/462	5	24/462	[ ]	80	368/462	8	36/462
	Stative	0	0/48	4	2/48	4	2/48	[ ]	79	38/48	13	6/48
	Range											

Table 4.16: Internal factors conditioning the use of all variants expressing necessity in 21st century QF (shading indicates favourable factors to variant choice).

With regards to *avoir à*, it is clear that inanimate (specifically impersonal) subjects favour its use, as do external sources of obligation and non-stative verbs. In HF, it was revealed above that use of this variant is largely restricted to one specific structural context. Based on the results in Table 4.17, it is clear that this pattern of use is also operative in QF, even more so in the case of the intervening direct object.

[112] Only marginals are provided for this factor group as there was a clear interaction with definiteness due to a lack of definite, inanimate tokens.

	HF		QF	
	%	**N**	**%**	**N**
Verbal object appears as intervening material	66	38/58	82	14/17
Complement of necessity verb *faire* in these constructions	58	22/38	57	8/14

Table 4.17: Rate of intervening direct object and *faire* in *avoir à* constructions in HF and QF.

Unlike the other variants, *devoir* does not seem to have a clear role in the QF system. Not only is it extremely infrequent (3%, n= 15), but no factors analyzed here favour its use, with the exception of animate subjects (though minimally). In HF, it is also relatively infrequent (10%, n= 77) and conditioned identically to *falloir que*, with the exception of animacy. No evidence was uncovered in either variety to support Larreya's (2004) claim that *devoir* is more 'insistent' than *falloir*. However, it has also been claimed in the literature that *devoir* is a formal variant (e.g., Lewis 2015), so it seemed appropriate to test this possibility.

The corpus used to represent 21st century QF in this study (*Le français en contexte: milieux scolaire et social*, Poplack & Bourdages 2008) has two built-in formality components that allow for the investigation into *devoir*'s purported association with style. This corpus was compiled with the intention of comparing the speech of students to that of their French teachers. These data offer a unique opportunity to observe whether there are differences in the speech of students, who have been shown elsewhere to be representative of the community norm (Poplack 2015) compared to that of French language-arts teachers, who are supposed to represent the standard. According to societal norms, the role of the French teacher is to not only to teach standard French, but also to transmit it in its 'correct' form. Studies using the data from this corpus have shown that the teachers' speech is indeed more formal than that of their students (Poplack 2015). The corpus also allows for the comparison of the speech of students in both informal (i.e., sociolinguistic interviews) and formal (i.e., year-end classroom debates) contexts.

In order to test whether *devoir* is associated with more formal styles of speech, teachers' rates of use of *devoir* was compared with that of their students. Using the results of an ongoing project at the Sociolinguistic Laboratory at the University of Ottawa, the use of *devoir* by students in formal and informal contexts is also considered. The results (Table 4.18) suggest that *devoir* does, in fact, appear to be a marker of formality.

QF teachers		**QF students (formal)**		**QF students (informal)**	
%	**N**	**%**	**N**	**%**	**N**
22	146/670	18	75/406	5	19/404

Table 4.18: Rate of *devoir* amongst students (in both formal and informal contexts) and teachers in 21st century QF.

It is used four times more frequently by the teachers than it is by their students in informal contexts and students use it more than three times as often in formal contexts than in informal ones. The function of *devoir* appears to be the expression of formal style, which supports the claims in the literature that it is stylistically conditioned (e.g., Lewis 2015). However, the lack of conditioning effects of the strength diagnostics contradicts the idea that *devoir* is the more 'insistent' variant (Larreya 2004) and its relative infrequency refutes the idea (e.g., Lewis 2015) that it is one of the most common variants.

4.4.1.3 Discussion

In this section, the expression of necessity in 21st century HF was compared to that of 21st century QF in order to identify if the same trends were operative in both varieties. While both varieties make use of the same inventory of variants to express necessity, their distribution differs. *Falloir que* dominates the QF system (80%), while in HF, it only accounts for only about one quarter (24%) of the variability with *falloir INF* representing the majority share (49%). Despite differences in variant distribution, the internal conditioning of each is nearly identical

across the two varieties and it is clear that each variant occupies its own niche in the system, although some of these distinctions are nearly lost in QF, where *falloir que* is used more than three quarters of the time.

Figure 4.3 summarizes the HF system. Both *avoir à* and *devoir* have specific roles, outlined above. HF speakers are left with three choices, which are first distinguished according to definiteness of the subject/referent: indefinite/generic contexts favour *falloir INF*, while definite/non-generic ones favour *falloir que* and *être obligé*. Speakers then must choose between these last two options and do so on the basis of various other factors which condition their use, including the source of obligation and the type of verb.

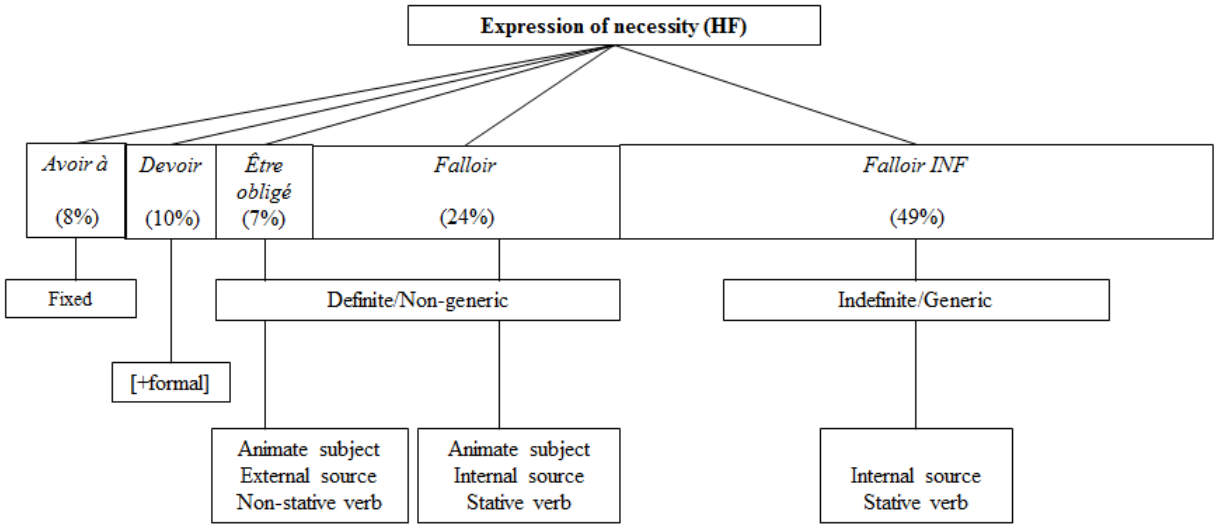


Figure 4.3: Summary of conditioning of variant choice in HF system of expressing necessity.

In the QF system (Figure 4.4), speakers mainly resort to *falloir que* as the default, but on the rare occasion that another variant is used, its conditioning matches nearly identically to that which was identified in HF, the only exception is that under *falloir INF*, source of obligation no longer plays a role. Because *falloir que* has become dominant in the QF system, it seems that it

has lost most of the conditioning that distinguished it from the other variants. The only factor that continues to minimally condition its use is the source of obligation, but this effect can be expected to eventually dissipate as well.

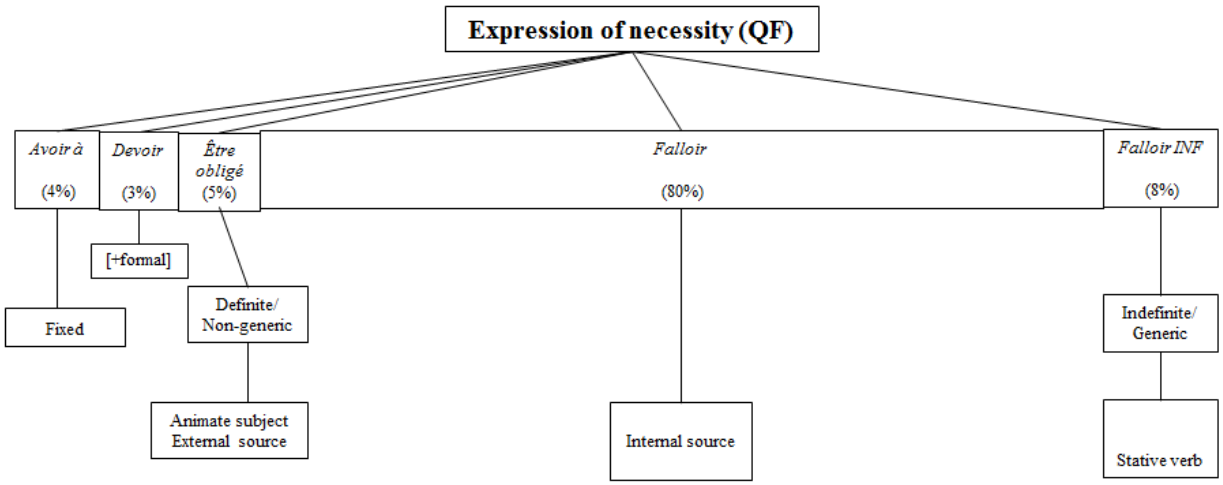


Figure 4.4: Summary of conditioning of variant choice in QF system of expressing necessity.

Since the main difference between the two varieties lies in their use of *falloir que* vs. *falloir INF*, a closer look at each of these two variants may help to explain where this divergence stems from. *Falloir que* is preferred with definite subjects, while *falloir INF* is preferred with indefinite/generic ones in HF. In QF, on the other hand, *falloir INF* is also favoured with indefinite/generic subjects (although used rarely), while *falloir que* is the main variant used with both definite and indefinite subjects.

One might suppose that the difference in the rates of *falloir* INF across the two varieties could have something to do with the proportion of indefinite/generic reference in each of the data sets under analysis. For example, if *falloir INF* is typically used to express indefinite/generic reference but the proportion of this variant is much lower in QF than in HF, then this could mean that in the QF data, there is simply less indefinite/generic reference. As shown in Table 4.19, this

assumption is false. The proportion of indefinite reference is nearly identical in each of the data sets.

	HF		QF	
	N	**%**	**N**	**%**
Total tokens with Indefinite/generic reference	453/802	56	292/510	57

Table 4.19: Amount of indefinite/generic reference in 21st century HF and QF.

If there is no difference is the proportion of indefinite/generic reference, then the difference must reside elsewhere, perhaps in the manner in which indefinite necessity is expressed. As shown Table 4.20, 74% of indefinite reference in expressions of necessity is expressed using *falloir INF* in HF.

	HF		QF	
	N	**%**	**N**	**%**
Falloir INF	336	74	31	11
Falloir que + indef. (pro)noun	35	8	232	79
Other variant + indef. (pro)noun	82	18	29	10
Total tokens with Indefinite/generic reference	**453**	**100**	**292**	**100**

Table 4.20: Display of the different methods employed in 21st century HF and QF to express indefinite/generic necessity.

However, in QF, the vast majority of indefinite/generic reference (79%) is expressed using *falloir que* with an overt indefinite subject. An example of each is given in (33) and (34). In other words, each variety employs a different method for expressing indefinite/generic necessity, which has contributed to proportional differences both here and in the use of the subjunctive (*falloir que* is a subjunctive-selecting context, while *falloir INF* is not).

(33) « C'était un examen qu'il ***fallait*** passer euh avant la licence. » (ESLO2/ZP063/335)

'It was an exam that you had to pass before getting Bachelor.'

(34) « Il ***faut*** tu_{indef} appelles peut-être quarante minutes à l'avance là. » (LFC[113]/152/393)

'You have to call maybe forty minutes ahead of time there.'

In the next section, the findings from 21st century HF and QF will be compared with the 20th century in order identify any recent changes in either system. The diachronic analysis may also shed light on this cross-variety synchronic divergence in the expression of indefinite/generic necessity.

4.4.2 Diachronic comparative analysis of the expression of necessity in HF and QF

The data analyzed in this section were subjected to both a real time and apparent time analysis (see Section 2.3 for a detailed description of these methods) in order to trace the diachronic trajectory of the expression of necessity in both varieties.

4.4.2.1 Overall rates

The results presented above revealed that in 21st century HF and QF, necessity is expressed using the same inventory of variants, with each filling clearly defined roles in the system. In order to identify possible change in progress, the 21st century findings are compared with the 20th century findings for both varieties.

In HF (Table 4.21), there are very few differences to note across this time period. *Avoir besoin* remains the least common variant, while the other minor variants (*devoir, avoir à, être obligé*) all hover around 10% in terms of overall usage. *Falloir que* remains stable, but there is a slight increase in the use of *falloir INF* between the 20th (44%) and 21st (49%) century, though this difference is not statistically significant, according to Fisher's exact test (p= 0.0509).

[113] This example is extracted verbatim from *Le français en contexte: milieux scolaire et social* (Poplack & Bourdages 2008), which is housed at the Sociolinguistics Laboratory at the University of Ottawa.

In QF (Table 4.21), most of the minor variants remain fairly stable across the time period examined with the exception of *être obligé* which dropped from 11% to 5%. However, there is a clear increase in the use of *falloir que* (from 55% to 80%) between the 20th and 21st century coupled with a clear decrease in the use of *falloir INF* (from 25% to 8%) over the same time period.

		20th		**21st**	
		N	**%**	**N**	**%**
HF	*falloir INF*	550	44	390	49
	falloir que	285	23	193	24
	devoir	151	12	77	10
	avoir besoin	28	2	21	3
	être obligé	104	8	59	7
	avoir à	126	10	63	8
	TOTAL	**1244**	**99**[114]	**802**	**101**[115]
QF[116]	*falloir INF*	192	25	42	8
	falloir que	429	55	406	80
	devoir	9	1	15	3
	avoir besoin	7	1	3	1
	être obligé	89	11	26	5
	avoir à	57	7	18	4
	TOTAL	**783**	**100**	**510**	**101**[117]

Table 4.21: Overall rates of use of each variant in the expression of necessity across time in HF and QF (20th and 21st centuries).

These results show that *falloir que* goes from being a majority variant (still being used twice as often as it is in HF) in the 20th century to nearly dominating the entire system by 21st century QF. The results for HF do not follow the same trajectory, at least within the time frame examined here.

[114] Due to rounding.
[115] Due to rounding.
[116] All results for 20th century QF are adapted from Lealess (2005).
[117] Due to rounding.

In an attempt to obtain more information on the trajectory of change in QF, each of the variants were extracted from *Les récits du français québécois d'autrefois*: *reflet du parler vernaculaire du XIX^e siècle* (Poplack & St-Amand 2009), which contains folktales and interviews with French rural speakers born between 1846 and 1895. As is illustrated in Table 4.22, there is little difference between the 19th and 20th century in terms of variant distribution, whereas between the 20th and 21st century, there is a clear change in the distribution of the main variants, with *falloir que* increasing, mostly at the expense of *falloir INF*[118].

	19th		**20th** [119]		**21st**	
	N	**%**	**N**	**%**	**N**	**%**
falloir INF	161	26	192	25	42	8
falloir que	311	50	429	55	406	80
devoir	30	5	9	1	15	3
avoir besoin	4	1	7	1	3	1
être obligé	62	10	89	11	26	5
avoir à	56	9	57	7	18	4
	624	**101**[120]	**783**	**100**	**510**	**101**[121]

Table 4.22: Overall rates of use of each variant in the expression of necessity across time in 19th, 20th, and 21st century QF.

This pattern is reminiscent of the one identified in Poplack, Lealess & Dion (2013) in their analysis of the subjunctive mood: the proportional importance *falloir que* within the subjunctive system increased across time, though the most drastic change was noted between the 19th and 20th centuries, not between the 20th and 21st centuries. This will be revisited in Chapter 5, where the relationship between the two changes is discussed in more detail.

[118] It should be noted that the 19th century corpus only contains old speakers, while the 21st century sample only contains young speakers; therefore, the increase in the use of *falloir que* between the 20th and 21st century may be inflated.
[119] Results for QF adapted from Lealess (2005:25).
[120] Due to rounding.
[121] Due to rounding.

4.4.2.2 External (social) factors

With regards to the effect of social factors on the expression of necessity in HF and QF, both sex and age were considered in order to assess their potential contribution to variant choice. Recall that in the case of change in progress, differences in the rates or internal conditioning[122] of young speakers (vs. older speakers) and of female (as opposed to male) speakers are likely. In fact, it is young females who are most likely to lead change in progress (Labov 2001). As is shown in Figure 4.5, there is no clear evidence of change in rate of variant use in apparent time (i.e., between different age cohorts) or in real time (i.e., comparing the different age groups across time in both the 20th and 21st centuries). In both time periods and in all age groups, *falloir INF* and *falloir que* represent the majority variants and remain fairly stable, hovering at around 45-50% and 20-25% respectively. The rest of the variants are used quite infrequently in each age cohort and across time.

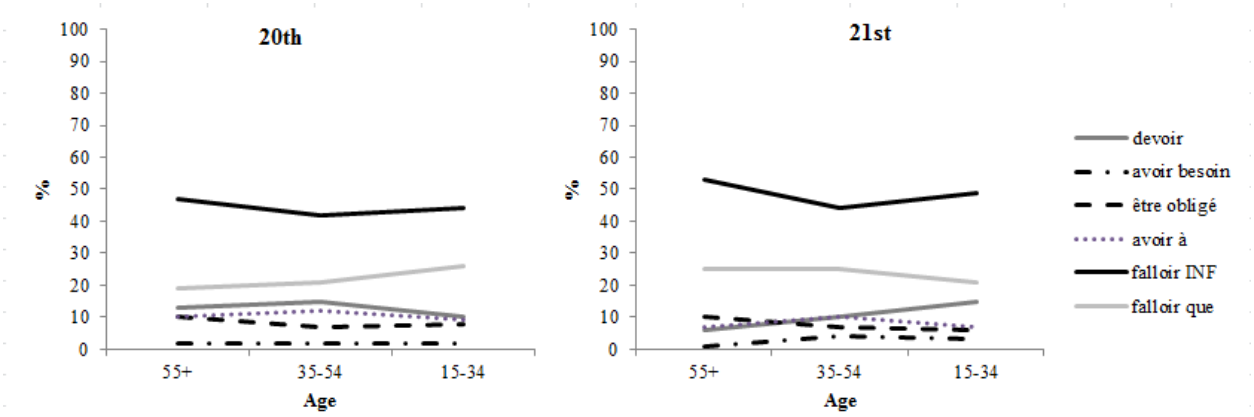


Figure 4.5: Variant choice in the expression of necessity according to age in 20th and 21st century HF.

[122] Internal factors for HF are not analyzed here because there is no evidence of change (based on rates). While there is evidence of change in QF, Lealess (2005) only analyzed the internal conditioning for some of the variants and most of the factor groups differed from those included here, so the analyses would not be easily comparable.

In Section 4.4.2.1, evidence of change in QF was presented, with *falloir que* increasing at the expense of the other variants. However, since the 21st century cohort only includes young speakers, it is possible that this is amplifying the change in progress, which is illustrated in Figure 4.6. It is clear that *falloir que* is the majority variant and is on the rise, while *falloir INF* has decreased substantially across time, the bulk of the change occurring between the 20th and 21st centuries in both cases.

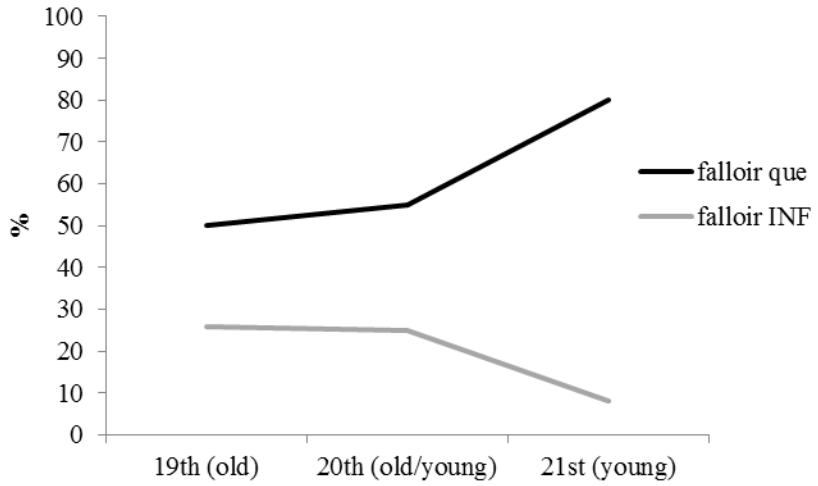


Figure 4.6: Choice between *falloir INF* and *falloir que* in 19th-21st century QF.

The results for variant choice according to speaker sex in both HF and QF[123] (Figure 4.7) reveal that there are no clear differences in either time period. In HF, *falloir INF* is the majority variant, followed by *falloir que*, and then by the others amongst both the male and female speakers. In QF, there are no clear differences in variant selection according to sex. However, the rate of *falloir que* is slightly more advanced amongst female speakers, which is to be expected in cases of change in progress (e.g., Labov 2001).

[123] Only results for 21st century QF are given here because this factor group was not tested in the analysis of the 20th century data (Lealess 2005).

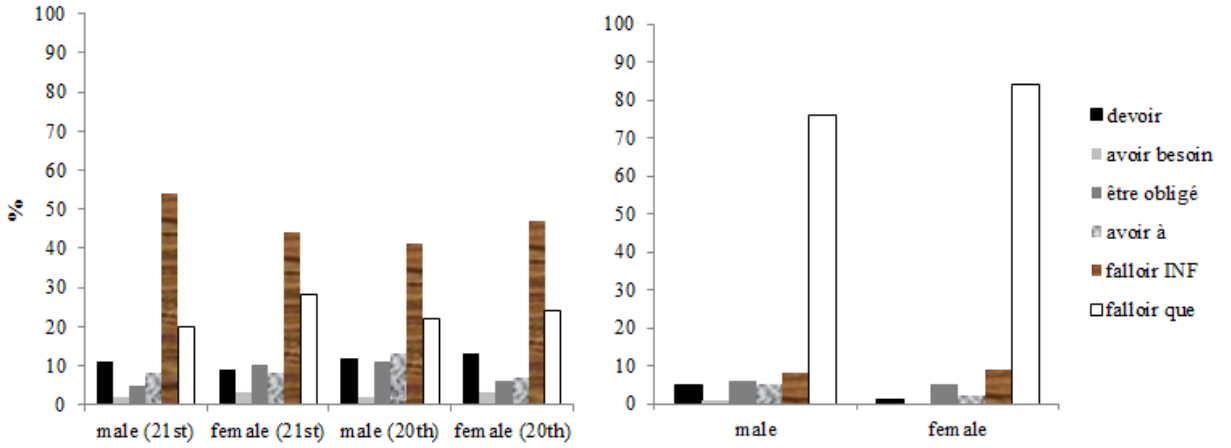


Figure 4.7: Variant choice in the expression of necessity according to age in 20th and 21st century HF and 21st century QF.

4.4.2.3 Discussion

The diachronic analysis of the expression of necessity undertaken here identified some clear differences across the two varieties in terms of the overall rate of the major variants (*falloir INF* and *falloir que*) (Figure 4.8).

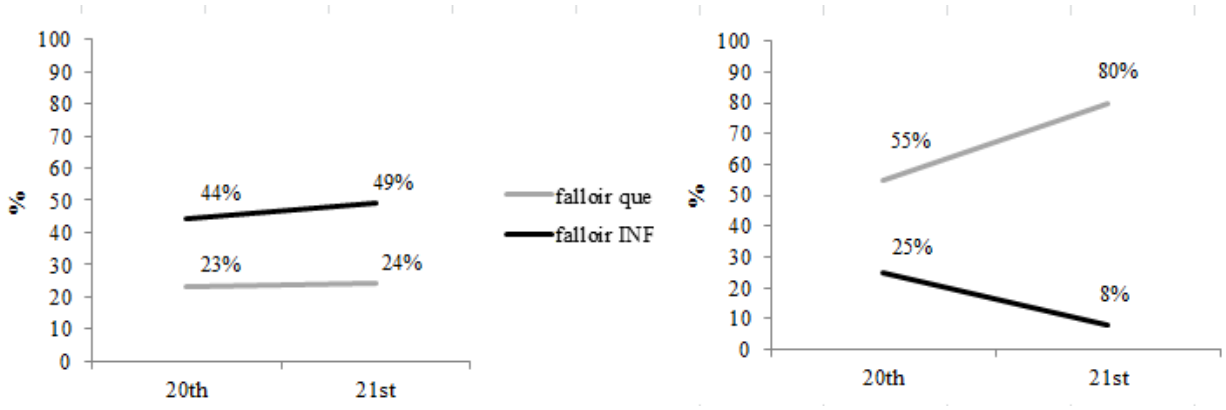


Figure 4.8: Overall rate of *falloir que* and *falloir INF* across time in HF and QF.

In HF, the majority variant is *falloir INF*, which slightly increases across the two time periods examined in this study (from 44% to 49%), while in QF, *falloir que* is by far the majority variant in both periods. In addition, the rate of *falloir que* increases drastically from 55% to 79% between 20th and 21st century QF, while *falloir INF* drops from 25% to 8% during the same

period of time. There is no notable change in the overall rate of *falloir que* in HF during this period. It was also noted in Section 4.4.1.2 that there were no distinct differences between the two varieties when it came to the rates or the internal conditioning of the variants in the 21st century cohorts, with the exception of *falloir que* and *falloir INF*, which are distinguished primarily by the definiteness of the subject or referent. HF speakers use *falloir INF* to express generic/indefinite necessity and *falloir que* to express definite necessity, while QF speakers mainly use *falloir que* for both purposes.

	Subject of complement of necessity	**HF 21st C**	**QF 21st C**
Falloir que	+2nd person (*tu/vous*)	1% (n=2)	79% (n=202)
	+*on*	1% (n=2)	2% (n=4)
	+Other 3rd (pro)noun	7% (n=25)	7% (n=18)
Falloir INF	Ø	92% (n=337)	12% (n=31)
	TOTAL	101% (n=366)[124]	100% (n=255)

Table 4.23: Display of the different methods employed in HF and QF to express indefinite/generic necessity with *falloir*.

In fact, once the distribution of the indefinite pronouns used in combination with *falloir que* in both varieties is broken down (Table 4.23), it is clear that the preferred method of expressing indefinite necessity for QF speakers is by use of the construction in (35), while in HF, it is overwhelmingly *falloir INF* that serves this purpose.

(35) *falloir que* + $tu_{indef.}$ + V_{subj}[125]

The distinction between *falloir INF* and *falloir que* is rarely commented on in prescriptive literature and when it is mentioned, there is no clear indication of which to use. For example, Grevisse and Goosse (2008:1469) state that the choice between the two is not restricted

[124] Due to rounding.
[125] As was shown in Chapter 3, the verb in (35) is nearly categorically conjugated with the subjunctive (where distinguishable).

and occurs freely. The findings of this study suggest that the choice of one over the other is, in fact, constrained by the definiteness of the subject, overt or otherwise. But why is there such a distinction? It is likely that the use of *falloir INF* to express indefinite/generic necessity stems from an earlier stage in the language. *Falloir INF* was first attested at the end of the 12th century, though it occurred in limited contexts and did not adopt its modern meaning until the 14th century, at which point it was attested with both infinitival and subjunctive complements (Rickard 1970:66). Its first appearance with a bare infinitive is argued to be "a continuous development from Latin" (Kjellman 1913:59; cited in Rickard 1970:69).

When *falloir* became associated with its modern meaning of necessity in the 14th century, two other impersonal variants were still being used in a layered system*: il estuet* and *il convient*, both of which could occur with an infinitive complement or a subjunctive one. Rickard (1970: 88) argues that by the time *falloir* was associated with its modern meaning, it inherited (in addition to the infinitive complement) the embedded subjunctive complement, which was already "well-established" with its predecessors. While the use of the infinitive with *falloir* can lead to ambiguity in terms of the referent, Rickard suggests that in cases where the context provides sufficient information as to the referent, it is an "admirably concise construction" (ibid., 86-90). He goes on to state that in cases where the referent is indefinite, this construction is fitting, although he does also mention the possibility of *falloir que* + indefinite *on*, noting, however, that it is only periodically attested in his analysis, which spans from the 14th to the 16th century. There is no mention of the use of indefinite *tu* with *falloir que*, as by most accounts (e.g., Blondeau 2008), this is fairly recent development in the language. Table 4.23 shows that *falloir que* + indefinite *on* continues to be rare in 21st century HF, while *falloir INF* retains its status as the preferred method of expressing indefinite/generic necessity.

To summarize, *falloir INF* predates *falloir que* and has long been attested to express indefinite necessity more so than *falloir que* + indefinite *on*. Therefore, it is unsurprising to find the same pattern of use in present day HF. It is clear that QF has changed, presumably from exhibiting a pattern of use similar to that found in HF, evidenced in part by the fact that in contexts in which *falloir INF* is still used, it is conditioned nearly identically across the two varieties, in particular in that it is favoured by indefinite reference. The question remains as to why QF has adopted *falloir que* as the majority variant to express both definite and indefinite necessity, and *falloir* que +tu_{indef} (which is largely absent from HF) for indefinite necessity in particular. This question inspired further analysis of the expression of pronominal indefiniteness more generally in both varieties.

4.4.3 Expressing pronominal indefiniteness in HF and QF

Contemporary French speakers can express indefiniteness by use of several different pronouns (*on, ils/elles, tu/vous*) or by use of a generic noun (e.g., *les gens, les femmes*, etc.). Since this analysis of the expression of necessity in HF and QF revealed that a clear preference for the form *falloir que* +tu_{indef} in the former, it seemed appropriate to further examine the role of *tu/vous* in expressing indefiniteness in contemporary French.

4.4.3.1 Indefinite *tu/vous* and *on*

Indefinite *tu/vous* is a fairly recent innovation that has begun to take over the role of indefinite *on*, first reported in Montreal French (Laberge 1977; Laberge & Sankoff 1980), and considered to be widespread in Canadian French (Posner 1997). It is speculated that one of the driving forces behind this shift may be to "relieve *on* of some of its burdens" (Ball 2000:67) since it can be used to express both definite and indefinite reference and thus can lead to

ambiguity. This is due, in part, to the fairly recent and ongoing use of *on* in place of *nous* for 1st person plural definite reference, a change that has nearly come to completion in QF (e.g., Blondeau 2003; Laberge 1977) . There are three different types of *on* that are available for use in French (Table 4.24). Here, the second type of *on,* which is in variation with *tu/vous* is of interest.

Definite *on*	**In variation with *nous* ('we')** e.g., « Alors que nous *on* peut faire un peu c'qu'on veut. » (ESLO2/YQ52/1853) 'Whereas us we can basically do what we want.'
Indefinite *on*	**In variation with *tu/vous* ('you', 'one')** e.g., « Avant tout ***on*** est soi-même et pas un Parisien pas un Marseillais. » (ESLO2/YQ52/4051) 'Before everything we are ourselves and not a Parisian or a Marseillais.'
Indefinite *on*	**In variation with *ils* ('they')** e.g., « Tous les jours ***on*** nous rappelle qu'y a l'bac. » (ESLO2/YQ52/782) 'Every day they remind us of the bac.'

Table 4.24: Different uses of *on* in contemporary French (adapted from Coveney (2003:174) and based on Laberge's (1977) classifications).

Quantitative studies of the use of indefinite *on* and *tu/vous* have been undertaken in both HF and QF and the general consensus is that the replacement of *on* by *tu/vous* is more advanced in QF than it is in HF. Deshaies (1985) found that indefinite *tu* occurred at a rate of 77% in Quebec City French and Laberge (1977:223,233) found a rate of 52% in Montreal French, with young male speakers using it more than the other cohorts. Studies of HF have found that the rate of indefinite *tu/vous* had reached 42% (mostly *vous*) by the 21st century in Tours (Ashby 1992)

and 48% in Picardy (Coveney 2003:175). While the HF rates are only slightly lower than what was found in Montreal by Laberge, they are quite a bit lower Deshaies' finding for Quebec City.

Comparing rates of indefinite *on* across studies can be problematic since it is often very difficult to determine whether it has an indefinite or definite reading. The ambiguous nature of *on* has been addressed in detail in Boutet (1986), Stewart (1995), and Laberge (1977). Laberge (1977) lists many of the possible clues that could lead to one reading over another but ultimately, most cases require contextual analysis to make a decision. This involves analysis of the speech leading up to the occurrence of *on* and figuring out if it refers to a specific group of people and whether or not it includes the speaker. Even when the analyst takes into account the context, many cases remain ambiguous. For example, Coveney (2003) excluded 964 cases of *on* in his study due to their ambiguous nature, while only retaining 539. Had those tokens been included as having a *possible* indefinite reading, it would have reduced the overall rate of indefinite *tu/vous* from 48% to 32%.

In order to investigate the link between expressing pronominal indefiniteness and expressing indefinite necessity in both HF and QF, a small subsample from each cohort represented in this study was analyzed. Six of the youngest speakers were selected from each corpus in order to be able to capture any potential recent innovations in the expression of indefiniteness. Each token of *on, tu,* or *vous*[126] was extracted. A total of 3957 tokens were retained for analysis and separated into definite and indefinite uses[127]. Despite the subjective nature of the coding process for indefinite *on* and the resulting lack of comparability across

[126] *Nous* was also extracted in order to compare with definite *on*.

[127] Each token was coded as having a clear definite reading or a possible indefinite/generic reading. This means that tokens that would have been coded as ambiguous in Coveney (2003) were included here as indefinite. This coding decision was made in order to ensure comparability with *falloir INF* tokens in the expression of necessity, which were also coded as having a clear definite referent or as having a possible indefinite/generic referent. Due to the subjective nature of the interpretation of these tokens, a subsample of these tokens underwent a reliability test in order to ensure accuracy.

studies, this study has the advantage of comparing four cohorts that were coded using the same protocol by a single analyst.

The results presented in Table 4.25 show that between indefinite *on* and *tu/vous*[128], the latter is by far the preferred variant in QF. It was already the majority variant in the 20th century (83%) but its use becomes nearly categorical by the 21st century (94%). Conversely, while there is an increase in the use of indefinite *tu/vous* by HF speakers across time (9% > 18%), the overall rate is *much* lower than what is found in QF.

	20th HF	**21st HF**	**20th QF**	**21st QF**
Indefinite *tu/vous*	9% n= 24/253	18% n=46/250	83% n= 914/1106	94% n=702/746

Table 4.25: Use of indefinite *tu/vous* in 20th and 21st century HF and QF.

As is shown in 4.26, it would appear that in QF, *on* has completely replaced *nous* as definite first person plural reference and is used only sparingly as an indefinite pronoun, while *tu* has largely taken over this role.

	20th HF	**21st HF**	**20th QF**	**21st QF**
Indefinite *on* (in variation with *tu/vous)*	91% n= 229/253	82% n= 204/250	17% n= 192/1106	6% n= 44/746
Definite *on* (in variation with *nous)*	65% n=186/284	100% n= 340/341	100% n= 551/552	100% n= 525/525

Table 4.26: Use of definite and indefinite *on* in 20th century HF and QF.

In HF, on the other hand, there is a clear increase in the use of definite *on* in place of *nous* between the 20th and 21st century coupled with a decrease in indefinite *on*, as it is slowly replaced by indefinite *tu*. This suggests that both varieties are undergoing a restructuring of the

[128] The vast majority of these tokens are *tu* in QF.

pronominal system and are changing in the same direction, but that QF is more advanced, at least in the case of *tu*.

4.4.3.2 The effect of pronominal indefiniteness on the expression of indefinite necessity

How do these results help to explain the differences in the expression of necessity between the two varieties? It is likely that an increase in indefinite *tu* precipitated the change in the expression of indefinite necessity in QF. The rise of *tu* in QF is not unique to this context: it has increased across time not only as an indefinite pronoun, but also as a discourse marker (e.g., *tu sais*) and as a definite 2nd person pronoun (instead of *vous*) (Thibault 1991a). Its use as an indefinite pronoun in QF rose most notably between the 20th and 21st centuries, evidenced by the fact that it was nearly absent from 19th century QF (Blondeau 2008). Thibault (1991a) suggests that such changes, including the use of *on* as a definite 1st person pronoun (in place of *nous*) are all part of the restructuring of the QF pronominal system. In other words, as *on* became increasingly associated with definite reference in place of *nous* in QF (as opposed to its original function as an indefinite pronoun), this resulted in an opportunity for other pronouns to take over the indefinite function of *on*, a role which was filled by *tu* (Thibault 1991a). While indefinite *tu* is used in HF, it does not occur at nearly the same rate, arguably due to the fact that *on* is still quite productive in indefinite contexts, perhaps because it only took over the role of the definite pronoun (in place of *nous*) more recently (Table 4.26).

Despite indefinite *on* still being widely used, *falloir INF* is the preferred method for expressing indefinite/generic necessity in HF over *falloir que* + indefinite *on*. It is likely that QF once exhibited the same pattern of use until the rise of *falloir que* + indefinite *tu*, which took over the role of *falloir INF* sometime between the 20th and 21st century. It should be noted that *falloir que* is also higher in 21st century QF definite contexts (80% vs. 46%; see Tables 4.10 and

4.16) than in HF, which indicates a general increase of the use of this construction in QF at the expense of the other variants.

4.5 Conclusion

In sum, this analysis of the expression of necessity in HF represents the first of its kind in that both the rates and conditioning of all of the variants expressing necessity in contemporary HF were examined using the variationist method. All previous quantitative studies in HF focus on exclusively on *falloir* and/or *devoir* (often labelled as the most frequent variants) and fail to account for the system as a whole. While this study has confirmed some of what has been suggested previously, including the association of *devoir* with formality and the proportional importance of *falloir* in the expression of necessity (e.g., Lewis 2015), it has challenged others (e.g., the frequency and insistent nature of *devoir,* Larreya 2004). Most importantly, it has revealed a clearly structured system of variation, with each variant occupying a clear role that remains stable within the time frame examined here. *Avoir à* is generally limited to a specific structural context, while *devoir* is largely stylistically conditioned. The remaining variants are distinguished according to the definiteness of the subject, with *falloir INF* favoured by indefinite subjects, and *falloir que* and *être obligé* favoured by definite subjects. These two are further distinguished through other internal factors, including verb type, and the source of obligation.

This synchronic comparative analysis revealed many similarities with the QF system, including the use of the same inventory of variants and as well as the fact that most variants occupy the same role in the system as they do in HF. A major exception is the prominence of *falloir que,* which has mostly replaced the other variants to the extent that only remnants of the original system remain, manifested in the internal conditioning of each variant. The main difference between the two varieties (i.e., the difference in rate with *falloir INF* and *falloir que*)

is at least partially explained by clear differences in the expression of indefinite/generic necessity in that HF speakers use *falloir INF*, while QF speakers mainly resort to *falloir que* + indefinite *tu*, which is likely due to the independent rise of indefinite *tu* in QF more generally.

In the next chapter, the findings presented in Chapter 3 and Chapter 4 are summarized, followed by a discussion of how and if the differences identified between the two varieties in both the subjunctive and expression of necessity systems are related to one another. The findings are then situated more generally within the context of the relationship between a transplanted dialect and the source variety.

Chapter 5 GENERAL DISCUSSION

5.1 Summary of findings

The purpose of this dissertation was to undertake a comparative variationist analysis of two linguistic variables in HF and QF. The intention was to provide some insight on the status of certain linguistic features in each variety and, in doing so, determine if the same variants exist in both varieties, and if so, confirm whether they behave similarly, both in terms of rates and linguistic conditioning (where possible). The systematic analysis of contemporary HF is crucial to our understanding of the characterization and development of QF since it is 1) the source variety, and 2) has not been subjected to the same level of contact with English as QF.

The first variable examined in the context of this study was the subjunctive- a variable that has been the object of countless studies in mainstream HF, but none within the variationist framework. The subjunctive was an ideal variable to study because it is thought to convey meaning and the fact that it does not in QF could be construed as loss (i.e., change). Results revealed, contrary to expectation, that the overall rate of subjunctive is lower in HF than in QF under both verbal and non-verbal governors. Further analysis revealed that these rate differences are actually not so much due to the propensity to express the subjunctive, but rather due to differences in the proportions of certain subjunctive governors and their idiosyncratic rates of subjunctive. It was also noted that, while most of the same governors and embedded verbs are as proportionally important in HF as in QF, there is no evidence of increased entrenchment across time between the 20th and 21st century in the former. Instead, subjunctive use in HF remains relatively stable within the time period examined in this study, both in terms of rates and

conditioning. However, it was inferred that HF must have undergone change at some point due to the fact that the strong associations of the subjunctive with certain governors was generally more advanced in 20th century HF than in was in 19th century QF. No evidence was found to suggest that any subjunctive replacement strategies (e.g., conditional, imperfect, future embedded verbs; infinitival complements) are at play, as suggested by some. Not only is the replacement of the subjunctive with infinitival complements limited to certain contexts, even where it is possible, it appears that infinitival complements may have been the preferred choice in certain contexts (e.g., with *falloir* to express indefinite/generic necessity) all along.

The internal conditioning of this variable was also examined in the present study in order to test for semantic and morphosyntactic motivations and several important findings emerged. In HF, just like in QF, all apparent semantic effects are, in fact, due to underlying lexical ones. In addition to lexical effects, structural factors also play a role in variant choice, most notably the presence of the complementizer *que*. Furthermore, the lack of semantic motivation for subjunctive use in HF represents evidence against the idea that it is only in QF that the subjunctive is meaningless as the result of change. On the basis of these findings, it can be concluded that both varieties changed in the same direction, or that the subjunctive was never semantically motivated in this context at all. In sum, while some of what had been alluded to in the literature (although often based on little or no evidence) is confirmed in this study, many of the claims, particularly those which suggest that the subjunctive is semantically conditioned and that it is in decline, were not borne out in this analysis.

The greatest difference found between the two varieties was the proportion of *falloir que* in each. In both HF and QF, *falloir que* is a high frequency governor and boasts nearly categorical rates of the subjunctive. However, the proportion of *falloir que* in QF accounts for

more than three quarters of the governor pool, while in HF, it only accounts for about one quarter. As such, it produced an apparent difference in the overall subjunctive association rates, which was eventually linked back to this fact. This led to a detailed investigation into the use of *falloir que* in both varieties.

Falloir que is often cited as one of the main variants used to express necessity in French (e.g., Lewis 2015). Its limited use in HF in the subjunctive variable context led to the obvious question: If HF speakers are not using *falloir que* to express necessity, what are they using? One of the main goals here was to answer this question by examining the second variable that forms the basis of this dissertation: the expression of necessity. The analysis of this variable (including all of the variants) is virtually absent from the HF literature. This study aimed to fill this gap by shedding light on the HF system in both the 20th and 21st centuries, and by comparing the patterns identified in HF to those found in QF.

The results revealed that in HF, the main variant used to express necessity is *falloir INF*, which represents nearly half of the data pool in both HF cohorts (44-49%). *Falloir que* is the next most frequent variant, but only accounts for about one quarter of the data (23-24%). The rest of the variants (*avoir besoin, être obligé, avoir à*, and *devoir*) each only account for small proportions of the data pool and each fills a clear role in the system, which remains stable across time. The two main variants (*falloir que/INF*) are distinguished primarily by the nature of the subject of the complement of necessity: *falloir INF* is favoured by indefinite ones and *falloir que* by definite ones.

In QF, on the other hand, *falloir que* is by far the preferred variant, increasing from 50% in the 19th century to 80% by the 21st century. While the internal conditioning of each variant closely resembles what was found in HF, the prominence of *falloir que* is such that speakers no

longer distinguish definiteness by the choice between *falloir que* and *falloir INF*. Rather, *falloir que* co-occurs with both definite and indefinite subjects equally and more frequently than in HF, the latter being expressed by use of indefinite *tu* (i.e., *falloir que* + tu_{indef}), which has undergone recent, rapid change in QF. Indefinite *tu* has completely taken over the role of indefinite *on* in QF, while in HF, *on* generally retains its indefinite function. With *falloir INF* representing the main variant to express indefinite necessity in HF (and presumably in QF at a prior stage), it is unsurprising that the rise in indefinite *tu* in QF would manifest itself in this context since *falloir que* + on_{indef} is exceedingly rare (and has been since the 14$^{\text{th}}$ century, see Section 4.4.2.3).

5.2 The (projected) trajectory of change

Three areas of the grammar in both HF and QF have been analyzed in the context of this study, and, while the two varieties exhibit a vast amount of similarities, three clear differences were identified: 1) there is a greater degree of entrenchment of the subjunctive with specific governors and embedded verbs in QF than in HF, 2) the necessity system in QF is dominated by *falloir que*, in contrast to the HF system, which exhibits a more nuanced system, and 3) there is an increase in the use of indefinite *tu* across time in QF in place of indefinite *on*, which is also underway in HF but at a slower pace. Importantly, little evidence of change was found in HF across the time period examined here which is in contrast to the situation in QF, where change was identified in each case, oftentimes as the result of comparison with 19$^{\text{th}}$ century data — a point of comparison that was not available for HF in the present analysis. How are these changes related, if at all?

A first assumption was that the limited proportional importance of *falloir que* in the subjunctive system in HF when compared with QF was due to differences in the expression of necessity, which is what led to the examination of this variable in the first place. However, as is

shown in Figure 5.1 and discussed in the earlier chapters, the most notable shift towards entrenchment of the subjunctive with *falloir* in QF took place between the 19th and 20th centuries and thus *precedes* the increase in the use of *falloir que* as a marker of necessity. In addition, the rise in the use of indefinite *tu* also underwent its most drastic change between the 19th and 20th centuries, also prior to the rise of *falloir que* in the expression of necessity, which took place between the 20th and 21st centuries. In other words, both the entrenchment of the subjunctive with *falloir que* and the drastic rise in indefinite *tu* occurred between 19th and 20th centuries in QF, while the most notable increase in the use of *falloir que* in the expression of necessity took place between the 20th and 21st century.

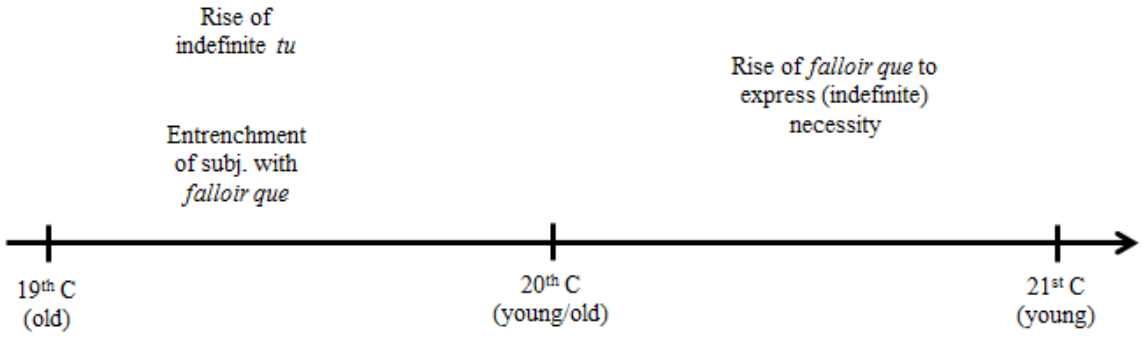


Figure 5.1: Timeline of change for three linguistic variables in QF.

While Poplack, Lealess & Dion (2013) identified a trajectory of change towards entrenchment of the subjunctive with *falloir que* between the 19th and 20th centuries (which remained fairly stable into the 21st century), the proportion of the data pool (i.e., frequency) occupied by *falloir que* did not increase across this period of time. In fact, it remained almost completely stable (19th: 59%, 20th: 65%, 21st: 63%, Poplack, Lealess & Dion 2013:166). In other

words, the frequency of *falloir que* did not increase across time in the subjunctive system, only its association with the subjunctive did. Why then is the overall increase in the use of *falloir que* to express necessity not reflected in the subjunctive system through an increase in its portion of the data pool? This is may be partially attributable to the fact that a large portion of the *falloir que* tokens included in the subjunctive analyses had to be excluded on the basis that they were ambiguous. In addition, the variable context for these two variables differs in that the proportion of *falloir que* in the subjunctive system is based on all of the potential governors which admitted a subjunctive in each time period, which, as discussed in Chapter 3, tend to fluctuate (especially in the case of infrequent ones). In other words, the portion of the governor pool accounted for by *falloir que* in the study of the subjunctive doesn't necessarily reflect its overall use in the language, which is more accurately captured in the variable context of the expression of necessity.

It is likely that the independent rise of indefinite *tu* (as a result of the restructuring of the pronominal system) in QF between the 19th and 20th centuries led to changes in expression of necessity, culminating in a clear preference for *falloir que* + tu_{indef} to express indefinite/generic necessity instead of *falloir INF*. Its rise in this context appears to be part of a more general pattern of strengthening of the *falloir que* construction in QF, which also manifests itself in definite contexts to a greater degree than in HF. It is not unlikely that advanced routinization of *falloir que* with the subjunctive in QF may also play a role in its increased use in the language more generally.

In HF, on the other hand, a much shorter period of time was available for analysis, which did not allow for comparison with the 19th century. As a result, little evidence of change was revealed, with the exception of a slight increase in the use of indefinite *tu* between the 20th and

21st centuries. The use of definite *on* (in place of *nous*) also increased, which suggests that like QF, HF is undergoing the same restructuring of the pronominal system, but at a slower rate.

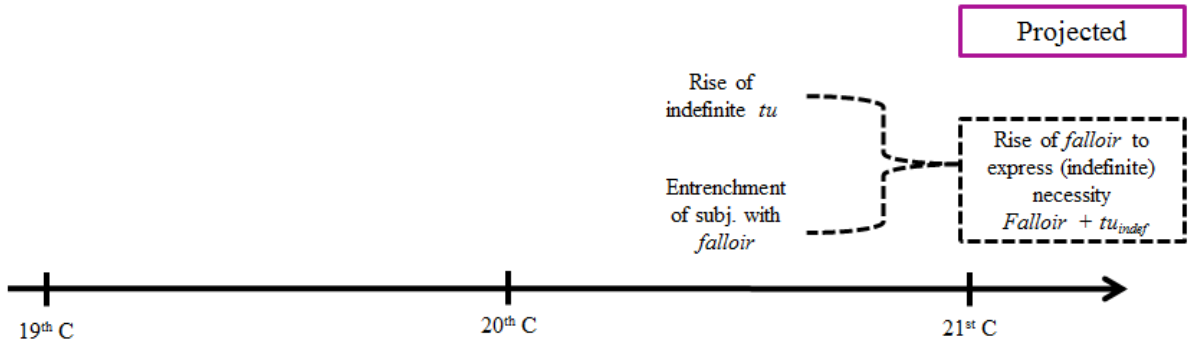


Figure 5.2: (Projected) trajectory of change in of three linguistic variables in HF.

With regards to the subjunctive, it was noted in Chapter 3 that it, too, has undergone change toward entrenchment, even if not shown within the time frame examined here. Based on a hypothetical earlier stage where subjunctive was fully productive in HF, as well as the fact that restrictions on its use are more advanced in 20th century HF than in 19th century QF, change toward entrenchment can be inferred[129]. Therefore, it appears that the subjunctive system in HF is undergoing the same trajectory of change as QF, but at a slower rate, as is illustrated in Figure 5.2. If the two changes suggested here as contributing to *falloir que*'s rise in the QF are lagging behind in HF, it is unsurprising to find no such change in HF's necessity system as of yet. However, it can be expected that it will follow the same path as QF. Analysis of pre-20th century HF would help to confirm the conclusions that have emerged from this study.

[129] One cannot know with any certainty whether or not the subjunctive was ever fully productive (semantically and/or morphosyntactically), or if this notion is just an artefact of prescriptive ideologies. However, these results suggest that subjunctive use in HF at an earlier stage was *more* productive than it is today.

5.3 The bigger picture

What do these findings reveal about the relationship between HF and QF and that between transplanted dialects and their source variety more generally? This study was undertaken in order to investigate perceived differences between QF and HF, specifically in the use of the subjunctive, which led to the analysis of the necessity system and the expression of pronominal indefiniteness. Two clear findings emerged from these analyses. First, both varieties appear to be moving in the same direction (albeit at different rates), with HF trailing QF in each case. Secondly, any noted differences between the two varieties were quantitative, not qualitative, in nature. In other words, in each case, the two varieties make use of the same variants and only differ slightly in their distribution. In fact, one of the only true differences between HF and QF emerged from the comparison of the expression of necessity in the 21st century, where *falloir que* now dominates the system in QF, showing that each variety favours a different lexical option from a shared pool of variants. However, this is not a deep grammatical difference. Rather, it is likely the result of two independent changes (the rise of indefinite *tu* and the entrenchment of the subjunctive with *falloir que*), which are also occurring in HF, but at a slower rate. This discovery is solely attributable to the application of the variationist comparative method, which provided the necessary methodological tools in order to go beyond rates of occurrence and proceed with a detailed analysis of patterns of use. Had rates of occurrence been the sole focus of this study, differences between the data sets would have been exaggerated and unexplained.

5.3.1 QF as a transplanted dialect

As was outlined in Section 1.2, transplanted dialects are commonly considered to exhibit certain characteristics that set them apart from the source variety, including the levelling and decomplexification of certain features, as well as the emergence of interdialect forms (e.g., Britain 2004). This study has revealed that QF does not fit this characterization at all, at least not when it comes to the variables examined here. There is no evidence of the use of interdialect forms (each variant identified in QF is also used in HF) or levelling (i.e., the replacement of irregular variants with regular, unmarked ones). The use of the subjunctive (i.e., the irregular variant) is actually on the rise in QF in specific structural and lexical contexts and HF is following the same trajectory of change. One could suggest that levelling may be at play in the replacement of indefinite *on* by indefinite *tu* or in the drastic rise of *falloir que* in the necessity system in QF. However, evidence of an increase in indefinite *tu* in HF would suggest if this is indeed the case, such patterns are not unique to QF.

Finally, the analysis of these variables has provided further support against the idea that QF is a simplified version of HF. While some may argue that the movement towards entrenchment of the subjunctive, the restructuring of the pronominal system, and the rise of *falloir que* to express necessity may, in fact, represent examples of decomplexification, it is important to note here that two of the three changes are also currently taking place in the source variety and that both varieties make use of the same variants in similar complex and structurally variable systems. If simplification is at play, it is the result of shared innovations, not differences stemming from separation. In fact, Lodge (2004) notes that the move toward increased regularity is a common trajectory of change not only in French, but in other Romance languages as well (Lodge 2004:222). The replacement of *nous* by *on* and the use of the periphrastic future figure

among his examples, both of which are occurring or have occurred in both HF and QF. Similarly, van der Molen (1923) attributes the decline of the subjunctive in French (which is refuted here) to a general pattern of simplification across Romance languages.

5.3.2 Influence of English?

While it is often speculated that differences between HF and QF are due to change in QF as a result of contact with English (e.g., Ager 1990), it has been repeatedly shown (e.g., Poplack & Levey 2010) that this is not the case. This study contributes to the debate by adding a horizontal benchmark of contemporary HF with which contemporary QF is compared. Such claims are not borne out in the comparative analyses conducted in the context of this dissertation. The general patterns and underlying grammars in the source variety, HF, (which has had little contact with English) and the transplanted variety, QF, (which has had ample contact with English) are nearly identical where comparison is possible, with the exception of some lexical differences.

The proportional differences in the use of certain lexical items (i.e., *falloir que* and indefinite *tu*) are also not attributable to contact with English. *Falloir que*'s potential English equivalents ('it is necessary' and 'have to') do not show a clear association with the subjunctive. The first 'it is necessary' does co-occur with the subjunctive in English, but is very rare, while 'have to' is not a subjunctive-selecting context at all (Kastronic & Poplack 2014). Indefinite *tu*, on the other hand, does have a clear equivalent in English (generic *you*); however, it also has an equivalent in other Romance languages (e.g., Spanish and Italian) (Coveney 2003:172). In addition, in Laberge's (1977) study of *on* vs *tu/vous*, she found that indefinite *tu* was frequent amongst working class males in Montreal who were unlikely to have had a great deal of contact with English at the time.

5.4 Future research

This comparative analysis of HF and QF revealed that, while both varieties surprisingly exhibit overwhelming similarities in three areas of the grammar, they nonetheless display clear idiosyncrasies in the form of lexical and proportional differences, which is likely due to the delayed trajectory of change in HF. It is likely that the changes noted in QF were embryonic at the time of split from HF and have simply progressed at a faster rate in the former. This finding raises an important question: why is HF lagging behind in change in progress, when the opposite scenario might have been expected (cf., colonial lag)[130]?

It is possible that dialect contact plays a role here. As discussed earlier, it is generally agreed upon that contact between different dialects played at least some role in the formation of contemporary QF, although the nature of this role is disputed. The HF benchmark examined here has also been subjected to dialect contact, resulting from increased mobility within France leading to a "proliferation of weak ties which expose the population to standardizing influences and facilitate the adoption of levelling changes" (Lodge 2004:226), although the influences reported here are far from 'standardizing'. Depending on the relative representation of each dialect and its majority variants, this could certainly have an effect on the trajectory and speed of linguistic change.

This dissertation represents an integral piece of a complex puzzle that will only be resolved with the analysis of more variables, both horizontally (as was done here) and vertically

[130] One reason to expect the opposite result is that in the only other comparative variationist study of QF and HF, Elsig & Poplack (2006) found that the interrogative system in HF had changed, while QF remained conservative, retaining an interrogative system reminiscent of an earlier stage, prior to colonization. It should be noted that the pattern of change identified in the present study is not new. For example, Trudgill (2004) noted parallel innovations in Southern hemisphere Englishes that were lagging or absent in the source variety. Similarly, American English is considered to be "the powerhouse for language change" Leech (2003:237), often surpassing its source (British English) in change in progress. Even in French, there are other documented cases of incipient change brought to Canada (e.g.,/t/, /d/ affrication, see Poirier 2008), which failed to progress in the same manner in the two varieties.

(by also examining a precolonial benchmark). Only then will a clear understanding of the true nature of the relationship between HF and QF emerge, which will, in turn, shed light on the connection between geographically distinct dialects and their source variety more generally.

REFERENCES

Abouda, Lotfi. 2002. Négation, interrogation et alternance indicatif–subjonctif. *Journal of French Language Studies* 12. 1–22.

Abouda, Lotfi & Olivier Baude. 2006. Constituer et exploiter un grand corpus oral: choix et enjeux théoriques. Le cas des ESLO. *Corpus en lettres et sciences sociales : des documents numériques à l'interprétation (Actes du colloque International d'Albi, juillet 2006)*, 143–150. Paris: Texto.

Ager, D. E. 1990. *Sociolinguistics and Contemporary French*. Cambridge ; New York: Cambridge University Press.

Alba-Salas, Josep. 2004. Lexically selected expletives: Evidence from Basque and Romance. *SKY Journal of Linguistics* 17. 35–100.

Anttila, Raimo. 1989. *Historical and Comparative Linguistics*. 2nd rev. ed. Amsterdam ; Philadelphia: John Benjamins.

Ashby, William J. 1976. The loss of the negative morpheme "ne" in Parisian French. *Lingua* 39. 119–137.

Ashby, William J. 1981. The loss of the negative particle "ne" in French: A syntactic change in progress. *Language* 57. 674–687.

Ashby, William J. 1992. The variable use of "on" versus "tu/vous" for indefinite reference in spoken French. *Journal of French Language Studies* 2. 135–157.

Ashby, William J. 2001. Un nouveau regard sur la chute du "ne" en français parlé tourangeau: s'agit-il d'un changement en cours? *Journal of French Language Studies* 11. 1–22.

Auger, Julie. 1990. *Les structures impersonnelles et l'alternance des modes en subordonnée dans le français parlé de Québec*. Québec: Centre international de recherche en aménagement linguistique.

Ayres-Bennett, Wendy, Janice Carruthers & Rosalind Temple. 2001. *Problems and Perspectives: Studies in the Modern French Language*. Harlow, England ; New York: Longman.

Baika, Gabriella. 1998. Deux verbes d'obligation en français: "devoir" and "falloir." *Studia universitatis Babes-Bolyai, Philologia* 43. 21–26.

Ball, Rodney. 2000. *Colloquial French Grammar: A Practical Guide*. (Blackwell Reference Grammars). Oxford, UK ; Malden, Mass., USA: Blackwell Publishers.

Barbaud, Philippe. 1994. Des patois au français : la catastrophe linguistique de la nouvelle France. In R. Mougeon & E. Beniak (eds.), *Les origines du français Québécois*, 79–100. Sainte-Foy: Les Presses de l'Université Laval.

Barral, Marcel. 1980. *L'imparfait du subjonctif: étude sur l'emploi et la concordance des temps du subjonctif*. Paris: Picard.

Bauche, Henri. 1929. *Le langage populaire*. Paris: Payot.

Benveniste, Emile. 1974. *Problèmes de linguistique générale*. Paris: Gallimard.

Bieler, A. 1972. L'enseignement du français aux États-Unis: étude sur le mode. *Etudes de linguistique appliquée* 17. 48–61.

Blanc, Michel. 1993. French in Canada. In Carol Sanders (ed.), *French Today: Language in its Social Context*, 239–256. Cambridge: Cambridge University Press.

Blanche-Benveniste, Claire, Mirielle Bilger, Christine Rouget & Karel van den Eynde. 1990. *Le français parlé: études grammaticales*. Paris: Éditions du Centre national de la recherche scientifique : Diffusion, Presses du CNRS.

Blanche-Benveniste, Claire & Colette Jeanjean. 1987. *Le français parlé: transcription et édition*. Paris: Didier Érudition.

Blanchet, Philippe & Nigel Armstrong. 2006. The sociolinguistic situation of contemporary dialects of French in France today: An overview of recent contributions on the dialectalisation of Standard French. *Journal of French Language Studies* 16. 251–275.

Blondeau, Hélène. 2003. The old "nous" and the new "nous": A comparison of 19th and 20th century spoken Quebec French. *University of Pennsylvania Working Papers in Linguistics* 9. 1–14.

Blondeau, Hélène. 2008. The dynamics of pronouns in the Québec languages in contact dynamics. In Miriam Meyerhoff & Naomi Nagy (eds.), *Social Lives in Language - Sociolinguistics and Multilingual Speech Communities. Celebrating the Work of Gillian Sankoff,* 249–271. Amsterdam: John Benjamins.

Börjeson, Lars. 1966. La fréquence du subjonctif dans les subordonnées complétives introduites par "que" étudiée dans des textes français contemporains. *Studia neophilologica* 38. 3–64.

Boutet, Josiane. 1986. La référence à la personne en français parlé : le cas de "on." *Langage et société* 38. 19–49.

Boysen, Gerhard. 1971. *Subjonctif et hiérarchie: étude sur l'emploi du subjonctif dans les propositions complétives objets de verbes en français moderne*. Odense: Univ. Press.

Branca-Rosoff, Sonia, Serge Fleury, Florence Lefeuvre & Mat Pires. 2012. Discours sur la ville. Présentation du corpus de français parlé parisien des années 2000 (CFPP2000). http://cfpp2000.univ-paris3.fr/CFPP2000.pdf.

Brennan, V. 1993. Root and Epistemic Modal Auxiliary Verbs. Amherst: University of Massachusetts Doctoral Dissertation.

Britain, David. 2004. Geolinguistic diffusion of language. In U. Ammon, N. Dittmar, K.J. Mattheier & P. Trudgill (eds.), *Sociolinguistics. An International Handbook of the Science of Language and Society*, 34–48. Berlin: Walter de Gruyter.

Brugmann, Karl. 1884. Zur Frage nach den Verwandtschaftsverhältnissen der indogermanischen Sprachen. *Internationale Zeitschrift für allgemeine Sprachwissenschaft* 1. 226–256.

Brunot, Ferdinand. 1922. *La pensée et la langue: méthode, principes et plan d'une théorie nouvelle du langage appliquée au français*. Paris: Masson et cie.

Bryant, William. 1980. Unequivocal passé composé/ imparfait. Contexts for "falloir" and "pouvoir." *The French Review* 53. 514–524.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Bybee, Joan & Sandra Thompson. 1997. Three frequency effects in syntax. *Annual Meeting of the Berkeley Linguistics Society* 23. 377–388.

Cajolet-Laganière, Hélène & Pierre Martel. 1995. *La qualité de la langue au Québec.* Quebec: Institut québécois de la recherche sur la culture.

Campbell, Lyle. 1998. *Historical Linguistics: An Introduction*. Cambridge, Mass: MIT Press.

Campbell, Lyle. 2013. *Historical linguistics: An introduction*. 2nd ed. Cambridge, Mass: MIT Press.

Chantefort, Pierre. 1976. Diglossie au Québec : limites et tendances actuelles. *Cahier de linguistique* 6. 23–53.

Charbonneau, Hubert. 1990. Le caractère français des pionniers de la vallée laurentienne. *Cahiers québécois de démographie* 19. 49–62.

Charbonneau, Hubert & André Guillemette. 1994. Les pionniers du Canada au XVIIe siècle. In Raymond Mougeon & Édouard Beniak (eds.), *Les origines du français québécois*, 59–78. Sainte-Foy: Les Presses de l'Université Laval.

Chaudenson, Hubert. 1994. Français en Amérique du Nord et créoles français : le français parlé par les immigrants du XVIIe siècle. In Raymond Mougeon & Édouard Beniak (eds.), *Les origines du français québécois*, 167–182. Sainte-Foy: Les Presses de l'Université Laval.

Chauveau, Jean-Paul. 1998. La disparition du subjonctif à Terre-Neuve, Saint-Pierre et Miquelon et en Bretagne: propagation ou récurrence. In P. Brasseur (ed.), *Français d'Amérique: Variation, créolisation, normalisation*, 105–119. Université d'Avignon, Centre d'études canadiennes.

Cohen, Marcel. 1965. *Le Subjonctif en français contemporain*. Société d'édition d'enseignement supérieur.

Comeau, Philip. 2011. A Window on the Past, a Move Toward the Future: Sociolinguistic and Formal Perspectives on Variation in Acadian French. Toronto: York University Doctoral Dissertation.

Copely, Bridget. 2006. What should "should" mean? Language Under Uncertainty Workshop, Kyoto University.

Corbeil, Jean-Claude. 2007. *L'embarras des langues origine, conception et évolution de la politique linguistique québécoise*. Montréal, Québec, Canada: Québec Amérique.

Côté, Marie-Hélène. 2012. Laurentian French (Québec): Extra vowels, missing schwas and surprising liaison consonants. In Randall Gess, Chantal Lyche & Trudel Meisenburg (eds.), *Phonological Variation in French: Illustrations From Three Continents*, 235–274. Amsterdam: John Benjamins.

Coveney, Aidan. 1996. *Variability in Spoken French: A Sociolinguistic Study of Interrogation and Negation*. Bristol: Elm Bank.

Coveney, Aidan. 2000. Vestiges of "nous" and the 1st person plural verb in informal spoken French. *Language Sciences* 22. 447–481.

Coveney, Aidan. 2001. *The Sounds of Contemporary French: Articlation and Diversity*. Exeter: Elm Bank.

Coveney, Aidan. 2003. "Anything you can do, tu can do better": "tu" and "vous" as substitutes for indefinite "on" in French. *Journal of Sociolinguistics* 7. 164–191.

Depraetere, Ilse & Susan Reed. Mood and modality in English. In Bas Aarts & April McMahon (eds.), *The Handbook of English Linguistics*, 269–290. Malden, MA, USA: Blackwell Publishing.

Deshaies, Denise. 1985. Références personnelles et types de discours en situation d'entrevue. In Pierre Léon & Paul Perron (eds.), *Le dialogue*, 77–91. Ottawa: Didier.

Dickinson, John Alexander & Brian J. Young. 2003. *A Short History of Quebec*. 3rd ed. Montreal ; Ithaca: McGill-Queen's University Press.

Dreer, Igor. 2007. *Expressing the Same by the Different: The Subjunctive vs. the Indicative in French*. Amsterdam ; Philadelphia: John Benjamins.

Dubois, Jean. 1969. Énoncé et énonciation. *Langages* 4. 100–110.

Dulong, Gaston. 1973. Histoire du francais en Amerique du Nord. In T.A. Sebeok (ed.), *Current Trends in Linguistics*, vol. 1. Paris: Mouton de Gruyter.

Durand, Jacques, Bernard Laks & Chantal Lyche. 2002. La phonologie du français contemporain: usages, variétés et structure. In C. Pusch & W. Raible (eds.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, 93–106. Tübingen: Gunter Narr Verlag.

Elsig, Martin & Shana Poplack. 2006. Transplanted dialects and language change: Question formation in Québec. *University of Pennsylvania Working Papers in Linguistics* 12. 77–90.

Fintel, Kai von. 2006. Modality and language. In Donald M. Borchert (ed.), *Encyclopedia of Philosophy – Second Edition*. Detroit: MacMillan Reference. http://web.mit.edu/fintel/modality.pdf.

Fintel, Kai von & Iatridou. 2008. How to say ought in foreign: The composition of weak necessity modals. In Jacqueline Guéron & Jacqueline Lecarme (eds.), *Time and Modality*. Dordrecht: Springer Netherlands.

Foulet, Lucien. 1919. *Petite syntaxe de l'ancien francais*. Paris: H. Champion.

Fox, Anthony. 1995. *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press.

Galliano, S., E. Geoffrois, D. Mostefa, K. Choukri & J.-F. Bonastre. 2005. ESTER phase II evaluation campaign for the rich transcription of French broadcast news. *Proceedings of Interspeech*. 1149–1152.

Glatigny, Michel. 1976. Remarques sur le subjonctif. *Le français dans le monde* 22. 17–25.

Godard, Daniele. 2012. Indicative and subjunctive mood in complement clauses: From formal semantics to grammar writing. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics 9*, 129–148. Paris: Colloque de syntaxe et sématique à Paris.

Google. 2016. Map of Ottawa-Hull. https://www.google.ca/maps/place/Hull,+Gatineau,+QC/@45.4511781,-75.8165227,12z/data=!3m1!4b1!4m5!3m4!1s0x4cce034729c5d671:0x6ed9ec6df9522dc0!8m2!3d45.428731!4d-75.7133658.

Gougenheim, G., R. Michéa, P. Rivenc & A. Sauvageot. 1964. *L'élaboration du français fondamental (ler degré): étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris: Didier.

Grevisse, Maurice & André Goosse. 1989. *Le bon usage: grammaire française ; Grevisse langue française*. 4 éd. Bruxelles: De Boeck-Duculot.

Grevisse, Maurice & André Goosse. 2008. *Le bon usage: grammaire française ; Grevisse langue française*. 14. éd. Bruxelles: De Boeck-Duculot.

Grevisse, Maurice & André Goosse. 2014. *Le bon usage: grammaire française ; Grevisse langue française*. 14. éd. Bruxelles: De Boeck-Duculot.

Grimm, David Ricky Leigh. 2015. Grammatical Variation and Change in Spoken Ontario French: The Subjunctive Mood and the Expression of Future Temporal Reference. Toronto: York University Doctoral Dissertation.

Gross, Maurice. 1975. *Méthodes en syntaxe: régime des constructions complétives*. (Actualités scientifiques et industrielles 1365). Paris: Hermann.

Gross, Maurice. 1978. Correspondance entre forme et sens à propos du subjonctif. *Langue française* 39. 49–65.

Guillaume, Gustave. 1929. *Temps et verbe*. Paris: Champion.

Guy, Gregory. 1988. Advanced varbrul analysis. In Kathleen Ferrara, Becky Brown, Keith Walters & John Baugh (eds.), *Linguistic Change and Contact*, 124–136. Austin, TX: Department of Linguistics, University of Texas at Austin.

Hacquard, Veronica. 2009. Modality. In C. Maienborn, K. von Heusinger & P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, 1484–1515. Berlin: Mouton de Gruyter.

Harris, Martin. 1978. *The Evolution of French Syntax: A Comparative Approach*. London: Longman.

Hockett, Charles F. 1950. Age-grading and linguistic continuity. *Language* 26. 449.

Hooper, Joan. 1975. On assertive predicates. In J. Kimbell (ed.), *Syntax and Semantics, Volume 4*, 91–124. New York: Academy Press.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Huot, Hélène. 1974. *Le verbe devoir: étude synchronique et diachronique*. Paris: Klincksieck.

Huot, Hélène. 1986. Le subjonctif dans les complétives : subjectivité et modalisation. In M. Ronat & D. Couquaux (eds.), *La grammaire modulaire*, 81–111. Paris: Minuit.

Imbs, Paul. 1953. *Le subjonctif en français moderne : essai de grammaire descriptive*. Strasbourg: Faculté des lettres de l'Université de Strasbourg.

INSEE (Institut national de la statistique et des études économiques). 2009. *Chiffres –clés sur une territoire : Commune d'Orléans*. http://www.insee.fr/fr/bases-de-donnees/esl/comparateur.asp?codgeo=com-45234.

INSEE (Institut national de la statistique et des études économiques). 2010. *Populations légales 2010 : 75-Paris*. http://www.insee.fr/fr/ppp/bases-de-donnees/recensement/populations-legales/departement.asp?dep=75.

Juneau, Marcel. 1972. *Contribution à l'histoire de la prononciation française au Québec: étude des graphies des documents d'archives*. Québec: Presses de l'Université Laval.

Juneau, Marcel. 1976. *La jument qui crotte de l'argent*. Québec: Presses de l'Université Laval.

Kastronic, Laura & Shana Poplack. 2014. The (North) American English mandative subjunctive in the 21st Century: Revival or remnant? *University of Pennsylvania Working Papers in Linguistics* 20. 70–80.

King, Ruth. 2000. *The Lexical Basis of Grammatical Borrowing: A Prince Edward Island French Case Study*. Amsterdam: John Benjamins.

Kjellman, Hilding. 1913. *La construction de l'infinitif dépendant d'une locution impersonelle en français des origines au XVe siècle*. Uppsala: Almqvist & Wiksell.

Kratzer, Angelika. 1991. Modality. In A. von Stechow & D. Wunderlich (eds.), *Semantics: An International Handbook of Contemporary Research,* 639–650. Berlin: Mouton de Gruyter.

La Presse. 1960. 1993. Montreal.

Labbé, Cyril & Dominique Labbé. 2013. La modalité verbale en français contemporain. Les hommes politiques et les autres. In Banks David (ed.), *La modalité, le mode et le texte spécialisé*, 33–61. L'Harmattan. (10 April, 2016).

Laberge, Suzanne. 1977. Étude de la variation des pronoms sujets définis et indéfinis dans le français parlé à Montréal. Montreal: University of Montreal Doctoral Dissertation.

Laberge, Suzanne & Gillian Sankoff. 1980. Anything you can do. In Gillian Sankoff (ed.), *The Social Life of Language*, 271–293. Philadelphia: Univ. of Pennsylvania Press.

Laboratoire ligérien de linguistique (Orléans). 1968. Enquêtes sociolinguistiques à Orléans. http://eslo.huma-num.fr/.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

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

Labov, William. 1982. Building on empirical foundations. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Current Issues in Linguistic Theory*, 17–92. Amsterdam: John Benjamins.

Labov, William. 1984. Field methods of the project on linguistic variation and change. In John Baugh & Joel Sherzer (eds.), *Language in Use: Readings in Sociolinguistics*, 28–54. Englewood Cliffs, NJ: Pretence-Hall.

Labov, William. 1994. *Principles of Linguistic Change. Vol. 1: Internal Factors*. Oxford: Blackwell.

Lagerqvist, Hans. 2003. L'autonomie du subjonctif français en proposition complétive. In Merete Birkelund, Gerhard Boysen & Poul Søren Kjærsgaard (eds.), *Aspects de la modalité*, 153–166. Berlin, Boston: De Gruyter.

Lalaire, Louis. 1998. *La variation modale dans les subordonnées à temps fini du français moderne: approche syntaxique*. Bern ; New York: P. Lang.

Larreya, Paul. 2004. L'expression de la modalité en français et en anglais (domaine verbal). *Revue belge de philologie et d'histoire* 82. 733–762.

Laurier, Michel. 1989. Le subjonctif dans le parler franco-ontarien : un mode en voie de disparition? In Raymond Mougeon & Édouard Beniak (eds.), *Le français canadien parlé hors Québec: aperçu sociolinguistique*, 105–126. Quebec: Presses de l'Université Laval.

Le Bidois, Georges & Robert Le Bidois. 1935. *Syntaxe du français moderne: ses fondements historiques et psychologiques*. Paris: Picard.

Lealess, Allison. 2005. "En français, il faut qu'on parle bien" : Assessing Native-like Proficiency in L2 French. Ottawa: University of Ottawa M.A. Mémoire.

Leech, Geoffrey. 2003. Modality on the move: The English modal auxiliaries 1961–1992. In R. Facchinetti, M. Krug & F. Palmer (eds.), *Modality in Contemporary English*, 223–240. Berlin: Mouton de Gruyter.

Lewis, Diana. 2015. A comparable corpus-based approach to the expression of obligation across English and French. *Nordic Journal of English Studies* 14. 152–173.

Lim, Esther. 1989. Emploi du subjonctif en français parlé. L.F.: U.P. Master's Thesis.

Lodge, R. Anthony. 2004. *A Sociolinguistic History of Parisian French*. Cambridge, UK ; New York, NY: Cambridge University Press.

Loengarov, Alexander. 2006. L'alternance indicatif/subjonctif dans les langues romanes. Motivation sémantico-pragmatique et grammaticalisation. Leuven, Belgium: Katholieke Universiteit Leuven Doctoral Dissertation.

Martin, R. 1983. *Pour une logique du sens*. Paris: PUF.

Martineau, France. 1985. L'élision variable de "que" dans le parler d'Ottawa-Hull. Ottawa: University of Ottawa M.A thesis.

Miller, Kimberely & Nathalie Dion. 2009. Measuring the Effectiveness of Proscription: Possessive Prespositions in French. NWAV 38, University of Ottawa.

Molen, William J. van der. 1923. *Le subjonctif: sa valeur psychologique et son emploi dans la langue parlée*. Zalt-Bommel: Van de Garde & Company's drukkerij.

Moreau, M.-L. 1986. Les Séquences préformées: entre les combinaisons libres et les idiomatismes. Le cas de la négation avec et sans "ne." *Le français moderne* 54. 137–160.

Morin, Yves-Charles. 1994. Les sources historiques de la prononciation du français au Québec. In Raymond Mougeon & Édouard Beniak (eds.), *Les origines du français Québécois*, 199–236. Sainte-Foy: Les Presses de l'Université Laval.

Mougeon, Françoise. 1998. *Quel français parler? Initiation au français parlé au Canada et en France*. Toronto: Éd. du Gref [u.a.].

Mougeon, Raymond & Édouard Beniak. 1991. *Linguistic Consequences of Language Contact and Restriction: The Case of French in Ontario.* Oxford: Clarendon Press.

Mougeon, Raymond & Édouard Beniak. 1994. Présentation. In Raymond Mougeon & Édouard Beniak (eds.), *Les origines du français Québécois*, 1–58. Sainte-Foy: Les Presses de l'Université de Laval.

Nerio, T. 1978. Problèmes et méthodes de la syntaxe du subjonctif français en français contemporain : ébauche pour une étude sociolinguistique. Tours: Université de Tours Doctoral Dissertation.

Nerio, T. 1979. Vers un modèle sociolinguistique de l'emploi du subjonctif en français contemporain. *Bulletin d'études de linguistique française Tokyo*. 14–40.

Nordahl, Helge. 1969. *Les systemes du subjonctif correlatif: étude sur l'emploi des modes dans la subordonnée complétive en français moderne*. Bergen-Oslo: Universitetsforlaget.

Palmer, Frank R. 2001. *Mood and Modality*. Cambridge, UK ; New York: Cambridge University Press.

Paolillo, John C. 2002. *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford, CA: CSLI Publications.

Parret, Herman. 1976. La pragmatique des modalités. *Langages* 10. 47–63.

Poirier, Claude. 1994. La langue parlée en Nouvelle-France : vers une convergence des explications. In Raymond Mougeon & Édouard Beniak (eds.), *Les origines du français Québécois*, 237–274. Sainte-Foy: Les Presses de l'Université Laval.

Poirier, Claude. 2008. L'assibilation des occlusives /t/ et /d/ au Québec. Le point sur la question. In Luc Baronian & France Martineau (eds.), *Le français d'un continent à l'autre. Hommages à Yves-Charles Morin*, 375–421. Quebec: Presses de l'Université Laval.

Poplack, Shana. 1989. The care and handling of a mega-corpus: The Ottawa-Hull French project. In Ralph W. Fasold & Deborah Schiffrin (eds.), *Current Issues in Linguistic Theory*, vol. 52, 411–451. Amsterdam: John Benjamins.

Poplack, Shana. 1990. Prescription, intuition et usage : le subjonctif français et la variabilité inhérente. *Langage et société* 54. 5–33.

Poplack, Shana. 1992. The inherent variability of the French subjunctive. In Christiane Laeufer & Terrell A. Morgan (eds.), *Theoretical Analyses in Romance Linguistics*, 235–263. Amsterdam: John Benjamins.

Poplack, Shana. 1997. The sociolinguistic dynamics of apparent convergence. In Gregory R. Guy, Crawford Feagin, Deborah Schiffrin & John Baugh (eds.), *Current Issues in Linguistic Theory*, vol. 128, 285–309. Amsterdam: John Benjamins.

Poplack, Shana. 2001. Variability, frequency, and productivity in the irrealis domain of French. In Joan L. Bybee & Paul J. Hopper (eds.), *Typological Studies in Language*, vol. 45, 405–428. Amsterdam: John Benjamins.

Poplack, Shana. 2015. Norme prescriptive, norme communautaire et variation diaphasique. In Kirsten Jeppesen Kragh & Jan Lindschouw (eds.), *Les variations diasystématiques et leurs interdépendances dans les langues romanes : actes du colloque DIA II à Copenhague*. Copenhague: Éditions de Linguistique et de Philologie.

Poplack, Shana & Johanne S. Bourdages. 2008. Le français en contexte: milieux scolaire et social. Research project funded by SSHRC (# 410-2005-2108 [2005-2008]).

Poplack, Shana & Nathalie Dion. 2009. Prescription vs. praxis: The evolution of future temporal reference in French. *Language* 85. 557–587.

Poplack, Shana, Allison Lealess & Nathalie Dion. 2013. The evolving grammar of the French subjunctive. *International Journal of Latin and Romance Linguistics* 25. 139–195.

Poplack, Shana & Stephen Levey. 2010. Contact-induced grammatical change: A cautionary tale. In Peter Auer & Jürgen Erich Schmidt (eds.), *Language and Space*, 391–419. Berlin, New York: Walter de Gruyter.

Poplack, Shana & Anne St-Amand. 2009. Les récits du français québécois d'autrefois : reflet du parler vernaculaire du 19e siècle. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 54. 511–546.

Poplack, Shana & Sali Tagliamonte. 2001. *African American English in the Diaspora*. Malden, MA: Blackwell.

Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane de Andrade Berlinck, Salvatore Digesto, Dora Lacasse & Jonathan Steuck. in press. Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. In Wendy Ayres-Bennett & Janice Carruthers (eds.), *Manuals in Linguistics: Romance Sociolinguistics*. Berlin, Boston: De Gruyter.

Poplack, Shana & Danielle Turpin. 1999. Does the futur have a future in (Canadian) French? *Probus* 11.

Poplack, Shana, Lauren Zentz & Nathalie Dion. 2012. Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition* 15. 203–225.

Portine, Henri. 1985. L'expression de la nécessité. *Revue internationale de philosophie* 39. 361–374.

Portner, Paul. 2009. *Modality*. Oxford ; New York: Oxford University Press.

Portner, Paul. 2011. Verbal mood. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, 1262–1291. Berlin: De Gruyter.

Posner, Rebecca. 1997. *Linguistic Change in French*. Oxford : New York: Clarendon Press ; Oxford University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London ; New York: Longman.

Rickard, Peter. 1970. "(il) estuet", "(il) convient", "(il) faut" and their contructions in Old and Middle French. *The French Language*. London: George G. Harrap & Co. Ltd.

Rivard, Adjutor. 1914. *Études sur les parlers de France au Canada*. Québec: J.-P. Garneau.

Roberts, Nicholas S. 2012. Future temporal reference in Hexagonal French. *University of Pennsylvania Working Papers in Linguistics* 18. 97–106.

Sand, Jørgen U. 1983. Le subjonctif en français oral. *Actes du VIIIe congrès des romanistes scandinaves*, 303–313. Odense University Press.

Sand, Jørgen U. 2003. Le subjonctif revisité. In Merete Birkelund, Gerhard Boysen & Poul Søren Kjærsgaard (eds.), *Aspects de la modalité*, 193–202. Berlin, Boston: De Gruyter.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. In Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, 140–161. Cambridge: Cambridge University Press.

Sankoff, Gillian & Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83. 160–188.

Sankoff, Gillian & Henrietta Cedergren. 1972. Sociolinguistic research on French in Montréal. *Language in Society* 1. 173–174.

Sankoff, Gillian & Suzanne Evans Wagner. 2006. Age-grading in retrograde movement: The inflected future in Montréal French. *University of Pennsylvania Working Papers in Linguistics* 12. 203–216.

Sankoff, Gillian & Diane Vincent. 1977. L'emploi productif de "ne" dans le français parlé à Montréal. *Le français moderne* 45. 243–256.

Sankoff, Gillian & Diane Vincent. 1980. The productive use of "ne" in Montreal French. In Gillian Sankoff (ed.), *The Social Life of Language*, 295–310. Philadelphia: Univ. of Pennsylvania Press.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Company.

Schleicher, August. 1861. *Compendium der vergleichenden grammatik der indogermanischen sprachen*. Weimar: Hermann Böhlau.

Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Hermann Böhlau.

Schreier, Daniel. 2003. *Isolation and Language Change Contemporary and Sociohistorical Evidence from Tristan da Cunha English*. Houndmills, Basingstoke, Hampshire; New York: Palgrave Macmillan.

Schuchardt, Hugo. 1868. *Vokalismus, Vol. 3*. Leipzig: Teubner.

Silva-Corvalán, Carmen. 2008. Limits of convergence in language contact. *Journal of Language Contact- THEMA 2*. 213–224.

St-Amand, Anne. 2002. Le subjonctif suivant une expression non-verbale. Ottawa: University of Ottawa Master's Thesis.

Stewart, Miranda M. 1995. Personally speaking … or not? The strategic value of "on" in face-to-face negotiation. *Journal of French Language Studies* 5. 203–223.

Sueur, Jean-Pierre. 1981. Les verbes modaux sont-ils ambigus? In J. David & G. Kleiber (eds.), *Recherches linguistics VIII. La notion sémantico-logique de modalité*, 165–182. Metz: Centre d'analyse syntaxique, Faculté des lettres et Sciences Humaines.

Tagliamonte, Sali. 2006. *Analysing Sociolinguistic Variation*. Cambridge, UK ; New York: Cambridge University Press.

Tagliamonte, Sali & Alexandra D'Arcy. 2007. The modals of obligation/necessity in the Canadian perspective. *English World-Wide* 28. 47–87.

Tagliamonte, Sali & Jennifier Smith. 2006. Layering, competition, and a twist of fate: Deontic modality in dialects of English. *Diachronica* 23. 341–380.

Thibault, Pierrette. 1991a. La langue en mouvement : simplification, régularisation, restructuration. *Linx* 25. 79–92.

Thibault, Pierrette. 1991b. Semantic overlaps of French modal expressions. *Language Variation and Change* 3. 191–222.

Thibault, Pierrette. 2001. Regard rétrospectif sur la sociolinguistique québécoise et canadienne. *Revue québécoise de linguistique* 30. 19–42.

Thibault, Pierrette & Michelle Daveluy. 1989. Some traces of the passage of time in Montreal French speech. *Language Variation and Change* 1. 19–45.

Thibault, Pierrette & Diane Vincent. 1990. *Un corpus de français parlé*. Quebec: CIRAL, Université Laval, Collection Recherches Sociolinguistique.

Thomason, Sarah Grey. 2001. *Language Contact: An Introduction*. Edinburgh: Edinburgh Univ. Press.

Togeby, Knud. 1966. La hiérarchie des emplois du subjonctif. *Langages* 1. 67–71.

Togeby, Knud. 1974. Emploi du temps aux dépens du mode. *Studia neophilologica* 46. 10–14.

Torreira, Francisco, Martine Adda-Decker & Mirjam Ernestus. 2010. The Nijmegen corpus of casual French. *Speech Communication* 52. 201–212.

Trudel, M. 1973. *La population du Canada en 1663*. Montreal: Fides.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: University Press.

Trudgill, Peter. 2004. *New-dialect Formation: The Inevitability of Colonial Englishes*. New York: Oxford University Press.

Trudgill, Peter. 2006. Dialect mixture versus monogenesis in colonial varieties: The inevitability of Canadian English? *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 51. 265–286.

Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society* 37. 241–254.

Valiquette, P. L. & R. Lesage. 1995. Sur quelques emplois de “devoir + infinitif” à l’imparfait. *Langues et linguistique* 21. 137–151.

Vandergheynst, Paul. 1982. L’emploi des modes et des temps dans les subordonnées introduites par “que”, observations sur un corpus de français parlé. Katholieke Universiteit te Leuven Master’s Thesis.

Villeneuve, Anne-José & Julie Auger. 2013. “Chtileu qu”i m’freumereu m’bouque i n’est point coér au monne’: Grammatical variation and diglossia in Picardie. *Journal of French Language Studies* 23. 109–133.

Villers, Marie-Eva de. 2008. Which Language for the Future? Inset 82: The Feminization of Titles in Québec. In Michel Plourde & Pierre Georgeault (eds.), *The French Language in Québec: 400 Years of History and Life Part Four II - French: A Shared Language*.

Vincent, Diane, Marty Laforest & Guylaine Martel. 1995. Le corpus de Montréal 1995. Adaptation de la méthodologie sociolinguistique pour l’analyse conversationnelle. *Dialangue* 6. 29–45.

Walker, James A. 2010. *Variation in Linguistic Systems*. New York, NY: Routledge.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. *Emprical Foundations for a Theory of Language Change*. Texas: University of Texas Press.

Wolfram, Walt & Natalie Schilling-Estes. 2003. Dialectology and linguistic diffusion. In Brian D. Joseph & Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, 713–735. Oxford, UK: Blackwell Publishing Ltd.

APPENDIX A

List of governors in 20th and 21st century HF

neg. = subjunctive occurred only under negated form of governor.

(neg.) = subjunctive occurred under both negative and affirmative form of governor.

Verbal governor	English translation
accepter neg.	to not accept
accorder	to grant
admettre	to admit
aimer (mieux/bien) (neg.)	to (not) like (better)
arranger	to arrange
arriver	to arrive
attendre	to wait
avoir (x) chance	to have (x) chance
avoir à coeur	to be committed
avoir aucun souvenir	to have no recollection of
avoir besoin	to need
avoir envie neg.	to not want
avoir le sentiment neg.	to not have the feeling
avoir l'impression (neg.)	to not have the impression
avoir marre	to be fed up
avoir pas de raisons	to have no reasons
avoir reussi à ce	to have succeeded at
avoir/être peur	to be scared
ce qui fait	which means
c'est pas	it's not
comprendre (neg.)	to (not) understand
connaître	to know
considérer neg.	to not consider
craindre	to fear
croire neg.	to not believe
demander	to ask
déplaire	to displease
dire	to say
envisager	to consider

espérer	to hope
s'étonner	to be surprised
être +attribute (neg.)	to (not) be +attribute
être/avoir moyen neg.	to not be/have a way
éviter	to avoid
exiger	to require
faire	to make
faire attention	to be careful
faire changement à ce	to make changes
faire du mal	to harm/hurt
faire en sorte	to ensure
faire plaisir	to please
faire rigoler	to make laugh
falloir (neg.)	to (not) be necessary/have to
gêner neg	to not bother/annoy
mériter	to earn/deserve
paraître impensable	to appear unthinkable
penser neg.	to not think
permettre	to permit/allow
plaire	to please
pouvoir dire neg.	to not be able to say
préférer	to prefer
promettre	to promise
regretter	to regret
rendre	to make
rester	the fact remains
se pouvoir	to be possible
se souvenir (neg.)	to (not) remember
sembler	to seem
souhaiter	to wish
suffire	to be enough
supporter neg.	to not support
supposer	to assume
surprendre	to surprise
tenir à ce que (neg.)	to (not) want
trouver bien	to find good
trouver chouettte	to find nice
trouver dommage	to find unfortunate
trouver dramatique	to find tragic

trouver insalubre	to find unsanitary
trouver logique	to find logical
trouver neg.	to not find
trouver pas mal	to find not bad
valoir le coup	to be worthwhile
valoir mieux	to be better
voir (neg.)	to (not) see
vouloir (neg.)	to (not) want
vouloir dire	to mean

Non-verbal governor	English translation
à condition que	provided/on condition that
à moins que	unless
à part que	except that
après que	after
au lieu que	instead
au/en cas que	in case that
d'autant plus que	especially
avant (même) que	before
bien que	even though
de manière à ce que	so that
des fois que	just in case
façon à ce que	so that
jusqu'à ce que	until
jusqu'à tant que	until
le fait que	the fact that
le temps que	the time that
maintenant que	now that
malgré que	despite
même que	even though
non pas que	not that
pour que	so that
pourvu que	as long as
quoi que	whatever
sans que	without

List of governors in 20th century QF (Poplack, Lealess & Dion 2013)

Verbal governor	English translation
accepter	to accept
admettre	to admit
aimer (mieux/bien)	to like
arranger	to arrange
arriver	to arrive
attendre	to wait
avoir hâte	to look forward to x
avoir/être peur	to be scared
comprendre (neg.)	to (not) understand
concevoir	to devise
considérer	to consider
croire (neg.)	to (not)believe
demander	to ask
désirer	to desire
dire	to say
empêcher	to prevent
entendre	to hear
espérer	to hope
être +attribute	to be +attibute
falloir	to be necessary/have to
forcer	to force
penser (neg.)	to (not) think
permettre	to permit/allow
préférer	to prefer
proposer	to offer
s'attendre	to expect
(s') étonner	to be surprised
se rappeler (neg.)	to (not) remember
se souvenir	to remember
souhaiter	to wish
surveiller	to monitor
tenir	to want
voir (veiller)	to watch
vouloir	to want

Non-verbal governor	English translation
à (en) ce que	that
à moins que	unless
au/en cas que	in case that
avant (même) que	before
en tout cas que	in case that
façon à ce que	so that
jusqu'à que	until
jusqu'à ce que	until
jusqu'à tant que	until
le temps que	the time that
mais que	when
pour que	so that
sans que	without
soit que	whether…or

APPENDIX B

Speaker name (CFPP corpus)	Initials
OZGUR KILIC	OZ
CHRISTOPHE ANDRE	CA
MARIE-ANNE ANDRE	MA
LAURENCE LEBLOND	LL1
STEPHANIE ZANOTTI	SZ
LUCIE DA SILVA	LD
PAULINE DE BORDES	PD
RAPHAEL LARIVIERE	RL
YVETTE AUDIN	YA
ROSEMOND EHRARD	RE
PATRICK BERNARD	PB1
ANITA MUSSO	AM1
MATHILDE LELONG	ML
MARIE LOUISE ORSIN	MO
BLANCHE DUCHEMIN	BD
REINE CERET	RC
JULIA TEIXEIRA	JT
KATIA TEIXEIRA	KT
PIERRE BEYSSON	PB2
MARIE BEYSSON	MB
BERNARD ROSIER	BR
MICHELINE ROSIER	MR2
THERESE LE VERN	TL
VALENTINE TESTANIER	VT
MATHIEU ROSIER	MR1
ELISA RYSNIK	ER
GABRIEL PUJADE	GP
NICOLE NOROY	NN
MIRA F	MF
PAUL SIMO	PS
PIERRE MARIE SIMO	PMS
GARY COLLARD	GC
JACQUELEINE PELLETIER	JP
MONIQUE CHASLON	MC
KILLIAN BELAMY	KB

LUCAS HERMANO	LH
LOUISE LIOTARD	LL2
JEANNE MALLET	JM
YOUCEF ZERARI	YZ
ABDEL HACHIM	AH
ANDRE MORANGE	AM2
MARIE-HELENE MATERA	MHM
YOUNES BELKACEM	YB
DOMINIQUE VALIN	DV
ISABELLE LEGRAND	IL
ANNI-LIES SIMO-GROEN	ALSG