Implicit norms and ‘school French’ forms

Linguistic cohesion of second-generation francophones in Victoria, BC

Suzanne Robillard

A thesis submitted to the
University of Ottawa
in partial fulfillment of the requirements for the
Doctorate in Philosophy degree in Linguistics

Department of Linguistics
Faculty of Arts
University of Ottawa

ABSTRACT

Within the francophone community of Victoria, British Columbia, first-generation (G_1) speakers provide conflicting usage models (Laurentian French and Hexagonal French) to second-generation (G_2) speakers. In addition, speaker interactions are greatly restricted due to severe community-level instability – a situation that seems likely to prevent G_2 speakers from converging on cohesive speech norms. However, long-term observation has uncovered the community-wide impression that G_2 speakers *do* share norms, but that their speech differs from both donor varieties. Instead, G_2 speech is thought to reflect a major influence from the context in which they primarily use French – the francophone school – and is thus impressionistically judged to have a rather 'standard' quality.

The central aims of this study are to 1) determine whether G_2 speakers do, in fact, share cohesive usage norms, and 2) substantiate the belief that G_2 French has been shaped by explicit instruction in prescriptive grammar ('school' French). I achieve these aims using the quantitative, empirical methods of COMPARATIVE VARIATIONIST ANALYSIS (Poplack and Tagliamonte 2001), an approach couched in VARIATION THEORY (Labov 1966, 1972, 1984). Using a newly constructed corpus of naturalistic G_1 and G_2 speech, I compare speakers' patterns for four linguistic variables (assibilation, *avoir/être*, future temporal reference, and yes/no questions) – each of which is diagnostic of differences between each of the donor varieties and the rules of school French – to the patterns attested in noncontact benchmark data.

Results for each of the variables independently show that G_2 speakers are cohesive in their treatment of three variables (assibilation, *avoir/être*, and future temporal reference), and are approaching cohesion in the fourth (yes/no questions), thus substantiating the hypothesis of G_2

linguistic cohesion. For all variables except yes/no questions, G_2 speakers are unidirectionally oriented towards LF patterns, but for yes/no questions they instead adopt more of a common target approach, which supplants the different 'default' interrogatives provided by the donor usage norms with *est-ce que*. As for whether G_2 speech has been influenced by explicit instruction in school French, the findings do not entirely bear this out. Although results indicate a superficial alignment with prescriptive injunction for both *avoir/être* and yes/no questions, this is not confined to G_2 speakers alone: G_1 speakers *also* use far more of the saliently privileged variants than would be expected from benchmark usage data. This suggests that the presence of these features at G_2 cannot be exclusively due to their exposure to prescriptive instruction. Supporting this conclusion is the fact that the rate of 'standard' variants does not consistently fluctuate according to degree of exposure to standard French.

These results have implications for the idea that restriction in contexts of usage is necessarily correlated with restriction in functional capacity as well (cf. Mougeon and Beniak 1991). Where it was possible to assess underlying conditioning, Victoria G_2 speakers demonstrated full alignment with benchmark Laurentian usage norms – demonstrating that they are not functionally restricted. This study also has implications for our understanding of how dialect contact might proceed in extreme minority situations. Despite acquiring language in a relatively 'unsettled' linguistic context, there was no evidence to suggest that G_2 speakers are subject to the processes associated with new-dialect formation or koineisation – with the possible exception of sociostylistic reallocation within the yes/no question system, which is not an outcome that is exclusive to situations of dialect contact. In other words, there is nothing to suggest that French in Victoria behaves all that differently from French in other communities, thereby disconfirming impressionistic assessments of G_2 linguistic peculiarity.

Pour ma grand-mère, Marie Robillard,

et les autres membres fondateurs de la communauté francophone de Victoria

ACKNOWLEDGEMENTS

I am sincerely grateful for the guidance, collaboration, and inspiration provided by my supervisor, Dr. Shana Poplack. Throughout my graduate studies, she has encouraged me to explore all angles of a problem, to develop and defend my ideas, and to deeply respect and honour the data we are fortunate enough to work with. It has been a privilege to have been trained by an internationally-renowned and respected variationist sociolinguist. In addition to her strong leadership in academic matters, her exemplary commitment to linguistic myth-busting has greatly inspired my own work in community outreach and popular linguistics. In innumerable interviews and public lectures, Shana has demonstrated how to rectify public misapprehensions with impeccable argumentation based on hard facts – all with customary grace and tact.

Indeed, the research done at the uOttawa Sociolinguistics Lab is ground-breaking, in that it refutes the erroneous beliefs that English irrevocably 'damages' French, or that 'good' French is whatever adheres to the rules put forth by the *Académie française*. In publicly responding to these ignorant and damaging attitudes, the Lab's findings have been central to the movement against linguistic insecurity that has recently gained both visibility and momentum in minority communities across the country. Although my initial motivation in undertaking doctoral work was simply to uncover the particulars about the French spoken in my own community, the Lab's lessons have encouraged and enabled me to provide community members (myself included) with the facts about our French, that we might direct our attention towards such known dangers as linguistic insecurity, rather than concentrate our efforts on non-issues to do with some subjective and flawed pursuit of language 'quality'.

Nathalie Dion – queen of the sub-analysis, fount of endless ideas and inspiration, and good friend – also deserves more thanks than I can put into words here. She provides unflagging support, kindness, and encouragement to every Labite that is fortunate enough to pass by her desk. Her own experiences as a minority francophone provided rich fuel for many discussions on language attitudes, linguistic insecurity, bilingual quirks, and more – often extending well past the Lab's closing time. My understanding of the link between Lab research and linguistic insecurity, as well as how to transmit that clearly to a general audience, would be woefully incomplete without Nath's insightful and astute observations.

The present study would not have been possible without the extensive support of Alex D'Arcy and her Lab crew at UVic, who were kind enough to host me as a Visiting Research Scholar in 2016-2017. They provided equipment and resources to complete the Victoria French Corpus, invited me to give and hear talks, and generally included me as one of their own during my tenure there. I am greatly indebted to Alex, as well as to Sonya Bird, for their enduring support and ongoing interest in my work on BC French – in a way, they started it all, because without their encouragement, I never would have even applied to grad school!

Over my years at uOttawa, I was privileged enough to benefit from financial support (including admissions scholarships from the university, an OGS scholarship, several bursaries, and – most notably – the André Lapierre scholarship in 2014) and intellectual support provided by my fellow Labites and colleagues from other labs. I am especially grateful to Félix, who has cheered me on and cheered me up throughout the most gruelling stages of this process. I also extend my sincere thanks to each of my uOttawa professors – Ana Arregui, Marc Brunelle, Stephen Levey, Éric Mathieu, Kevin McMullin, Shana Poplack, Andrés Salanova, and Jamie White – for their contributions to my academic success via lessons and seminars, comp

supervisions, and research collaborations. The members of my evaluation committee, Stephen Levey, Richard Clément, André Lapierre, and Doug Walker, also have my deepest gratitude for their insightful, stimulating, and thorough reports, which have immeasurably improved the final product (of course, any remaining errors or inconsistencies are my own). Thank you also to Donna, for her help in ensuring continual deliveries of the dissertation in progress and for being there early every morning to open the Lab for me, and to Suzanne and Roxanne for their patient answers to my many administrative questions.

There are countless other friends who have my thankfulness for their contributions to the preservation of my health and happiness throughout this process: ma chère marraine Suzanne (the absolute *best* Ottawa host anyone could ask for), the Spoons gang, Capital Drift drivers and staff, the Granger Zone, the Gals, and all the rest. Thank you for reminding me that there is life outside of school and for giving me various outlets to switch off my work-brain. Thank you also to my family and to Britt, for being by my side through it all with words of encouragement (sometimes gentle, sometimes firm, always needed) – I quite literally could not have done this without your endless support. My cat, Loki, also deserves some recognition, since he regularly dragged me away from my desk for more treats, playtime, and pats. Though most unhelpful progress-wise, the benefits to my state of mind were always extremely welcome. Lastly, to my dearest Serge, who has soldiered through many a challenge with a smile on his face and good humour in his heart, your perseverance continues to be a true inspiration. Thank you for being by my side through it all.

TABLE OF CONTENTS

Chapter 1. Rationale, review of the literature, and research aims 1

1.1. General introduction 1

1.2. Dialect contact in context 11

1.2.1. Dialect contact within established communities 11

1.2.2. Dialect contact in new territory 16

1.2.3. Dialect contact within the Victoria francophone community 21

1.3. Research questions and theoretical framework 28

1.4. Roadmap 34

Chapter 2. The Victoria francophone community 36

2.1. Establishment and permanence 36

2.2. Francophone domains 38

2.3. Contact between HF- and LF-origin individuals 44

2.4. G_2 French in the school 49

2.5. Summary 53

Chapter 3. Data and method 55

3.1. Construction of the *Victoria French Corpus* (VFC) 57

3.1.1. The sociolinguistic interview 59

3.1.2. Sampling 62

3.1.3. Cohorts 65

3.2. Comparative benchmarks 69

3.3. Diagnostic variables 74

3.3.1. Assibilation of coronal plosives 79

3.3.2. Avoir/être 80

3.3.3. Future temporal reference 81

3.3.4. Yes/no questions 83

3.4. Extralinguistic factors 85

3.4.1. Contextual restriction 87

3.4.2. Contact with individuals and with infrastructure 94

3.4.3. Familiarity with school French 96

3.4.4. Other characteristics 98

3.5. Summary 101

Chapter 4. Assibilation as a window onto competing usage norms 103

4.1. Background 105

4.2. Benchmarks 107

4.2.1. Implicit usage norm in LF 107

4.2.2. Implicit usage norm in HF 108

4.3. Variable context and exclusions 111

4.4. Results 113

4.4.1. Overall 113

4.4.2. HF_1 and HF_2 cohorts 114

4.4.3. Extralinguistic factors 116

4.5. Summary and interim conclusions 123

Chapter 5. *Avoir/être*: A salient conflict between usage and the standard 125

5.1. Background 127

5.2. Benchmarks 130

5.2.1. Normative French 130

5.2.2. Implicit usage norm(s) 132

5.3. Variable context and exclusions 136

5.4. Hypotheses about variant selection 139

5.5. Results 145

5.5.1. Overall....145
5.5.2. Implicit usage norms for all cohorts....147
5.5.3. Extralinguistic factors....153
5.6. Summary and interim conclusions....157

Chapter 6. Future temporal reference: A three-way conflict with non-salient variability 161

6.1. Background....162
6.2. Benchmarks....163
6.2.1. Normative French....163
6.2.2. Implicit usage norm in LF....165
6.2.3. Implicit usage norm in HF....168
6.3. Variable context and exclusions....169
6.4. Hypotheses about variant selection....171
6.4.1. Polarity....172
6.4.2. Temporal distance....173
6.4.3. Adverbial specification....175
6.4.4. Style (subject type)....176
6.5. Results....178
6.5.1. Overall....178
6.5.2. Comparison of internal conditioning for LF_1 and LF_2 cohorts....181
6.5.3. Comparison of internal conditioning for HF_1 and HF_2 cohorts....184
6.5.4. Extralinguistic factors....188
6.6. Summary and interim conclusions....192

Chapter 7. Yes/no questions: A three-way conflict with salient variability....194

7.1. Background....195
7.2. Benchmarks....198
7.2.1. Normative French....198

7.2.2. Implicit usage norm in LF 206

7.2.3. Implicit usage norm in HF 208

7.3. Variable context and exclusions 211

7.4. Hypotheses about variant selection 218

7.4.1. Subject type 219

7.4.2. Factor groups related to lexical identity of verb 221

7.4.3. Style (topic) 222

7.5. Results 224

7.5.1. Overall 224

7.5.2. Comparison of internal conditioning for LF_1 and LF_2 cohorts 227

7.5.3. Comparison of internal conditioning for HF_1 and HF_2 cohorts 235

7.5.4. Extralinguistic factors 238

7.6. Summary and interim conclusions 246

Chapter 8. Finding the big picture: Summary and general discussion 249

8.1. Review of study aims and methods 249

8.2. Main findings 251

8.3. Connections to the dialect contact literature 255

8.4. Final reflections 261

LIST OF TABLES

Table 1.1. All diagnostics used to evaluate G_2 speech 32

Table 2.1. Origins of French-speaking international immigrants within the BC francophone population [adapted from Chavez & Bouchard-Coulombe (2011), Table 3.7.2.2] 45

Table 3.1. Familial configurations of G_2 speakers 66

Table 3.2. Number of participants, by cohort 69

Table 3.3. Full list of contexts in which French is claimed to be used 89

Table 3.4. Restriction index scores, by cohort 89

Table 3.5. Language use in different contexts, by individual 91

Table 3.6. Contact with infrastructure and with individuals, by individual 95

Table 3.7. Contact with francophone infrastructure, by cohort 95

Table 3.8. Contact with francophone individuals, by cohort 96

Table 3.9. Exposure to school French, by cohort 98

Table 3.10. Proportion of HF-origin individuals in network, by cohort 98

Table 3.11. Socioeconomic class (average house price in neighbourhood), by cohort 99

Table 3.12. Proportion of life spent in Victoria, by cohort 100

Table 4.1. N tokens of assibilation in each cohort 112

Table 4.2. Overall %assibilation for each cohort 113

Table 4.3. Rate of lenited variants in voiceless (/t/) and voiced (/d/) contexts 115

Table 4.4. Cross-tabulation (2x2 contingency table) of generation and assibilation variants 117

Table 4.5. Cross-tabulation (2x2 contingency table) of origins and assibilation variants 117

Table 4.6. Rates of each variant for non-categorical speakers 118

Table 4.7. Bivariate analysis of extralinguistic effects on assibilation selection for the HF_2 cohort 120

Table 4.8. Bivariate analysis of effect of %HF in network on assibilation selection for the LF_2 cohort 121

Table 5.1. Broad topics broached during sociolinguistic interviews 145

Table 5.2. Overall %AVOIR for each cohort 146

Table 5.3. Overall %AVOIR in previous studies 146

Table 5.4. %AVOIR and %POOL by lexical verb (all cohorts aggregated) 148

Table 5.5. Selection of *avoir* according to lexical verb, by cohort 152

Table 5.6. Effect of speech style on selection of *avoir* 153

Table 5.7. Cross-tabulation (2x2 contingency table) of generation and *avoir/être* variants 154

Table 5.8. Cross-tabulation (2x2 contingency table) of origins and *avoir/être* variants 154

Table 5.9. Bivariate analysis of extralinguistic effects on *avoir* selection for the LF$_2$ cohort 156

Table 6.1. Overall rate of synthetic future for each cohort 178

Table 6.2. Overall rate of synthetic future in previous studies 179

Table 6.3. Multivariate analysis of synthetic future selection for the LF$_1$ cohort 182

Table 6.4. Multivariate analysis of synthetic future selection for the HF$_1$ cohort 185

Table 6.5. Cross-tabulation (2x2 contingency table) of generation and future temporal reference variants 189

Table 6.6. Cross-tabulation (2x2 contingency table) of origins and future temporal reference variants 189

Table 6.7. Bivariate analysis of extralinguistic effects on synthetic future selection for the HF$_1$ cohort 190

Table 6.8. Bivariate analysis of extralinguistic effects on synthetic future selection for the LF$_1$ cohort 191

Table 7.1. Summary of diagnostics used for yes/no questions 211

Table 7.2. Effect of affirmative (+) and negative (–) polarity on yes/no question variant selection ... 217

Table 7.3. Overall distribution of yes/no question variants for each cohort 225

Table 7.4. Overall distribution of yes/no question variants in previous studies 225

Table 7.5. Internal conditioning of yes/no questions for the LF$_1$ cohort 228

Table 7.6. Internal conditioning of yes/no questions for the LF$_2$ cohort 233

Table 7.7. Internal conditioning of yes/no questions for the HF$_1$ cohort 236

Table 7.8. Internal conditioning of yes/no questions for the HF$_2$ cohort 238

Table 7.9. Cross-tabulation (2x2 contingency table) of origins and yes/no question variants 239

Table 7.10. Cross-tabulation (2x2 contingency table) of generation and yes/no question variants ... 240

Table 7.11. Bivariate analysis of extralinguistic effects on yes/no question variant selection for the LF_1 cohort 241

Table 7.12. Bivariate analysis of extralinguistic effects on yes/no question variant selection for the LF_2 cohort 244

Table 7.13. Bivariate analysis of extralinguistic effects on yes/no question variant selection for the HF_2 cohort 245

LIST OF FIGURES

Figure 2.1. Origins of provincial immigrants who claim French as their mother tongue within the BC francophone population [adapted from Chavez & Bouchard-Coloumbe (2011), Table 3.7.1] 45

Figure 4.1. Distribution of each variant across individuals in the HF_2 cohort 116

Figure 5.1. Rate of *avoir* according to lexical verb in benchmark LF data 141

Figure 5.2. Rates of benchmark *avoir* usage in benchmark LF data and in Victoria data 150

Figure 7.1. Stylistic expectations for yes/no questions in school French and in the implicit usage norms of HF and LF 224

Chapter 1. Rationale, review of the literature, and research aims

1.1. General introduction

The francophone community in Victoria, British Columbia[1] differs from many other minority francophone communities in Western Canada in that it cannot directly trace its origins back to the 19th century diaspora of Quebecois settlers (a trajectory described in more detail by (Allaire 2014), but is instead primarily composed of relatively recent arrivals from regions such as Quebec, Ontario, or France (Chavez and Bouchard-Coulombe 2011; see also Chapter 2). Whereas the French spoken in long-established Western francophone enclaves is thought to have directly evolved from Quebec French and to have preserved its features to a large extent (Walker and Canac-Marquis 2016), it is not clear whether the same could be said of Victoria, given the historical impermanence of French in the region and the diversity of French dialects it is presently home to. In other words, there is nothing to suggest – on the face of things – that the presence of French in Victoria is sustained through what might be thought of as 'normal' community transmission, whereby a single, uniform, and stable sociolinguistic model is transmitted directly from adults to children over the course of many generations.

In such a 'normal' scenario, very young children initially acquire the uniform structural base – the 'accent' (phonetic inventory, phonological rules), lexical expressions, and morphosyntax – that characterises their parents' speech, as well as the implicit rules that dictate which forms

[1] 'Victoria' in this dissertation should be taken to mean the municipalities of the Greater Victoria Area, as well as those of the West Shore and the Saanich Peninsula.

might be (dis)favoured by a given linguistic context. By age seven or so, they have also very readily acquired systematic norms for the social or stylistic stratification of variability; that is, the intuitive knowledge that certain forms are attributable to members of upper/lower social classes or perhaps to more formal/informal registers (Labov 1964). To the best of our collective knowledge, parents do not need to facilitate language transmission in any special way: under normal conditions, transmission of both linguistic forms and their associated (grammatical or sociostylistic) functions is effectively an automatic process.

In communities where only one model is provided (i.e., where both the uniform structural base and the sociostylistic conditioning are shared across all the members of a community, often termed a SPEECH COMMUNITY[2]), there are generally no *major* deviations from this model in the immediately subsequent generation. This is not to say that the linguistic behaviour in such communities is fixed or static, nor that every pattern is always acquired uniformly by all individuals within a community (cf. Smith, Durham, and Richards [2013]; though see also Guy [1980] and Walker and Meyerhoff [2013]). Once children begin to interact with individuals outside their immediate household, they begin to speak differently from their parents – though

[2] For a much more complete description of this, see (Labov 1966; 2001), whose conception of the speech community – while not uncontested (cf. Patrick 2004; Romaine 1982) – often factors into the key assumptions of many community-based sociolinguistics studies. In minority communities particularly, this notion may be somewhat problematic, since communities that form clear social or cultural groups (e.g. 'individuals identifying with francophone culture') may nonetheless display *dis*continuities in linguistic usage (cf. Mougeon and Nadasdi 1998). One purpose of this dissertation is to determine whether there is, in fact, linguistic unity within the Victoria francophone population, and if so, the extent that this is shared across all sectors of the community. It can therefore speak directly to the question of discontinuity once the results have been presented, and so further discussion of this issue is saved for the concluding chapter.

they still maintain essentially the same structural base (Hazen 2006). The interaction with individuals within a community – particularly with peers – is what is thought to move internally-motivated language change forward (cf. Labov 1966; 2001; Payne 1980), but in general, any such changes within a speech community proceed quite incrementally (though, of course, successive generations moving a change in the same direction might create a large difference between the oldest generation and the youngest in a given sample).

The Victoria francophone community does not seem like one of these 'typical' situations – at least on the surface of things. Instead, it bears more of a resemblance to situations described in the literature on dialect contact, where children acquiring language are exposed to a number of different dialects, and must consequently choose from this assortment of linguistic options to form their own patterns for speech. The source of these different dialects is the new arrivals, here termed first-generation (G_1) individuals, who in Victoria form the bulk of the francophone population. They arrive equipped with the IMPLICIT USAGE NORMS[3] that characterise their language or dialect of origin, which differ from the established local norms to a greater or lesser degree in terms of their structural base (i.e., lexicon, phonological inventory or rules, grammatical categories, and so on) and their implicit or explicit sociostylistic associations with certain forms (Labov 2001, 87). Upon establishing themselves in the new locale, G_1 individuals may engage in ACCOMMODATION to their interlocutors (Giles et al. 1987), which can have the

[3] The notion of 'implicit usage norms', a key tenet of variationist sociolinguistics, refers to the underlying patterns that probabilistically govern the selection of one variant over another when there are several variants that share the same referential meaning (cf. Weinreich et al.'s [1968, 100] STRUCTURED HETEROGENEITY). In this dissertation, 'implicit usage norms' should be interpreted as interchangeable with the terms 'sociolinguistic model' and 'patterns for speech/speech patterns'.

effect of shifting their implicit norms to some degree, perhaps by causing them to converge[4] with the norms of their interlocutors, but most G_1 speakers are limited in the amount of vernacular reorganisation they can perform (Labov 2001, 415–17) and therefore largely maintain the norms they arrived with. The children of these individuals (second-generation, or G_2, speakers), whose language is developing within the bounds of the new community, are therefore provided with at least two distinct usage models for speech: the parental model, as well as the model that is locally predominant.

The literature on dialect contact, critically reviewed in the coming sections, suggests that the outcome of many such scenarios is G_2 cohesion. That is, even in situations where G_1 speakers maintain their dialectal differences, G_2 speakers will nonetheless manage to converge on shared and cohesive implicit norms for speech. Furthermore, many studies find that G_2 speakers achieve cohesion under a range of more or less challenging dialect contact conditions; for instance, when there is a pre-established local dialect to conform to, but also when dialects of the language in question are not yet established (e.g. during colonisation of new territory). In some cases, G_2 speakers may align themselves with a locally predominant dialect. Alternatively, they may innovate novel patterns that consist of a 'mixture' of the source dialects or use forms that are

[4] Accommodations that make the linguistic behaviour of two interlocutors more similar can be labelled with the term CONVERGENCE (as opposed to DIVERGENCE, which acts to emphasize the linguistic differences between speakers). For some scholars, 'convergence' designates the outcome of successive, community-wide accommodations that occur over the long term and have resulted in permanent or lasting changes to the dialects or languages in question – something which can only be inferred once they have already occurred (e.g. Auer and Hinskens 1996). For others, it simply indicates "the enhancement of inherent structural similarities found between two linguistic systems" (Bullock and Toribio 2004, 91), which can be assessed synchronically, and which does not necessarily imply that permanent or lasting changes have taken place. My use of the term in this dissertation conforms to the latter definition.

'intermediate' between the two source dialects – though it should be noted that these qualifications are not unproblematic, something which is discussed further in Section 1.2.

The Victoria situation adds to the body of work on dialect contact in that it takes place under conditions whose effect on linguistic outcomes at G_2 has not yet been directly studied. In Victoria, in addition to having to grapple with exposure to multiple sociolinguistic models over the course of their linguistic development, children acquiring French as a first language face the extra complications of doing so in an environment that is characterized by francophone population mobility[5], assimilation to English (the majority language by far, as the francophone community only makes up approximately 1% of the Greater Victoria population; cf. Chavez & Bouchard-Coulombe, [2011]), and contextual restriction of French to a limited range of social domains. Despite a long history of francophone settlement in the region (from at least 1843 [Agnew et al. 1987]; see also Section 2.1), the community is persistent but very unstable. It is in a constant cycle of loss (from assimilation or emigration) and renewal (from a steady stream of new arrivals from various regions), with the result that although the community never quite dies out, its members are ever-changing (Chavez and Bouchard-Coulombe 2011). Rates of migration (both incoming and outgoing) for francophones in British Columbia are among the highest in the country. It is contextually restricted – being a minority community, the use of the French dialects in contact is confined to one or two contexts only. In addition, the prevalence of English acts to

[5] As an anecdotal case in point, my younger brother and I had almost completely different sets of teachers from kindergarten to 12 – as well as different administrative and support personnel – despite only being four grades apart. More recent personal communications with teachers at the francophone school reveal that the vast majority of the current K-12 faculty has only been in place for a few years at the very most, and that the bulk of these new teachers have only resided in Victoria for a short time. This highlights the speed of francophone population turnover in Victoria.

depress the general frequency of interactions in French, as many francophones incorporate anglophone individuals into their social networks, or (especially at G_2, since these individuals are all bilingual) opt to use English in encounters with members of the francophone community. Consequently, there are very few opportunities for regular exposure to or interaction in French.

Furthermore, the social contexts in which G_1 and G_2 speakers use French are different. Broadly speaking, for G_2 speakers as a whole, French is used primarily in the classroom context and, to a lesser extent, in the home and/or in community activities; for G_1 speakers, French is used primarily in the home context, with family and friends who remained in the region of origin, and, more rarely, in community activities. The connection of G_2 speakers to Victoria's only francophone school, *École Victor-Brodeur*, is noteworthy, since this institution is arguably among the most important contributors to community cohesion. Schools play a pivotal role in French language transmission and maintenance more generally (e.g. in many minority francophone communities across Canada), as well as in the development of francophone identity and community solidarity. In Canada, parents whose first language is French have the constitutional right to send their children to francophone school, and many Victoria parents avail themselves of this right (see Section 2.4). All classes – save for second-language instruction in English or Spanish, for example – are taught in French, and students are required to speak French on school grounds. This rule is diligently enforced by school staff – although it is frequently contravened when students are not directly under the eye of a teacher – with the result that French tends to be used primarily in the classroom. This is not to say that *no* G_2 individuals prefer French for peer socialisation, only that such individuals tend to be in the minority and that many peer-level interactions (e.g. at recess, in the school hallways) take place in English. Opportunities for face-to-face interaction in French are thus severely limited – even within some

francophone households, for instance, where G_1 parents may be assimilating to English alongside their children. These circumstances (described in more detail in Chapter 2) seem likely to be a major barrier to G_2 acquisition and subsequent maintenance of French in the first place – let alone achieving shared and cohesive implicit usage norms that distinguish their French from that of their parents.

Nonetheless, however unlikely it may seem at the outset, long-term participant observation of this community has revealed that many G_2 speakers are not only fully native and fluent speakers of French (see Chapter 2; also Section 3.4), they are also identified as speaking a variety of French that is particular to them (or at least different from other varieties of French), and as having a strong sense of shared, francophone identity ('*franco-colombien.ne*'). Indeed, many community members, as in (1)–(5) below, claim that the French spoken by the younger generation is both characteristic of G_2 speakers specifically, as well as somehow different from that of G_1 speakers (e.g., more 'English', more 'international', more 'standard', a 'mix' of influences, or even '*Brodeurien*' – a reference to the name of the local francophone school).

(1) Il y a une influence– un– une influence anglaise assez forte, uh, et puis il a des– des influences un peu plus internationales, mais un peu français, un peu québécois– bien beaucoup québécois je dirais, mais ça peut être des franco-canadiens de l'Ontario […] Donc c'est ça, il y a beaucoup d'influences et ça fait un, um, méli-mélo de français (*inc.*) un peu *Brodeurien*. [VFC.HF$_2$.Frédéric.00/57/57][6]

'There's an influence– a– a fairly strong English influence, uh, and there are some some– influences that are a bit more international, but a bit French, a bit Quebecois– well a lot Quebecois I would say, but it might be from French-Canadians from Ontario. […] So that's it, there are many influences and it creates a, um, hodgepodge of Frenches (inc.) a bit Brodeurien.'

(2) T'es plus– on est plus habitués à entendre plein de différents accents, je trouve qu'ici– que beaucoup d'autres endroits. C'est comme, il a pas tout w– c'est un mélange de québécois et de français de France. [VFC.HF$_2$.Hugo.00/22/57]

'You're more– we're more used to hearing plenty of different accents, I find than here– than in many other places. It's like, there isn't all w– it's a mix of Quebecois and French from France.'

(3) C'est juste comme ça qu'ils [étudiants à Brodeur] parlent parce que t'as un mélange de cultures et puis forcément quand t'as un mélange de cultures tu vas avoir– comme, tout est mélangé. [VFC.HF$_2$.Annique.01/18/45]

'That's just how they [Brodeur students] talk because you have a mix of cultures and inevitably when you have a mix of cultures you'll have– like, everything is mixed.'

[6] Codes refer to corpus name, cohort, speaker pseudonym, and timestamp of utterance. All utterances are transcribed verbatim from interview recordings.

(4) C'est assez international, des fois, les étudiants à Brodeur, parce que c'est vrai qu'ils parlent– ils parlent des français de partout dans le monde. [VFC.HF_2.Frédéric.00/48/28]

'It's fairly international, sometimes, the students at Brodeur, because it's true that they speak– they speak Frenches from all over the world.'

(5) Donc ils [G_2 speakers] ont– ils ont un français plutôt– un français plus standard, pas pi– pas typiquement québécois. [VFC.HF_1.Hélène.00/42/41]

'So they [G_2 speakers] have– they have a French that's rather– a more standard French, not pi– not typically Quebecois.'

In fact, speaking anecdotally as a third-generation member of this community myself, observations of the type in (1)–(5) strike me as very commonplace, and the informal realisation that outgroup individuals (including G_1 speakers) perceive our French as somehow unique is what prompted my investigation of this community in the first place – especially given the highly unusual (and rather adverse) conditions in which it would have been acquired.

The frequent mention of the francophone school as a cultural linchpin in the Victoria francophone community suggests that it may provide a solution to the challenge of converging on a cohesive set of norms for French in a highly restricted dialect contact environment. In addition to French instruction, the provincial francophone school board mandates its schools to "develop students' francophone cultural identity and sense of belonging," with staff acting as "cultural transmitters" (*Politique H-800-1* [Conseil scolaire francophone de la Colombie-Britannique 2007]; translations mine throughout), alongside their more academic or pedagogical responsibilities. The school effectively acts as a major community hub: it offers its own daycare and preschool service, hosts francophone events, clubs, and extracurricular programs after school

hours, and boasts its own community theatre and outdoor recreation facilities. Beyond education in the minority language, these amenities act to encourage the formation and maintenance of G_2 francophone social networks outside the classroom. The school acts to unify G_2 individuals under a single cultural umbrella (which might push speakers towards linguistic unity as well), and also provides speakers with a daily context in which to hear and use French. It therefore seems entirely possible that the influential role played by the school in shaping francophone identity or in providing cultural touchstones might directly translate into a similarly important influence over G_2 linguistic structure.

Indeed, if one goal of the dissertation is to explore how (or if) cohesive implicit usage norms are acquired by G_2 speakers in the first place, another is to determine *which* of the available models in their environment has exerted the greatest effect on the resulting linguistic norms – something which is discussed further in Section 1.2.3. The multitude of different characterisations that surfaced in community evaluations (cf. (1)–(5); 'mixed', 'standard', '*Brodeurien*', and so on) shows that, while G_2 speakers are identified as distinct from the rest of the community somehow, nailing down exactly *what* makes them different is perhaps not so straightforward. On the one hand, it might be the case that G_2 speech patterns *are* uniform, either because they have all converged on a single model that is present in their environment, or because they all incorporate the same elements from a number of different sources (two possibilities I explore further below). On the other hand, it is possible that G_2 speakers are *not* linguistically uniform at all, and that the various impressionistic assessments more or less correspond to actual linguistic differences between speakers. For example, *some* G_2 speakers may be more 'standard', whereas others are 'mixed', others are 'Laurentian', etc. Finally, the variety of descriptions could merely be a flaw of impressionistic assessments – that is, G_2 speakers might

be different from G_1 speakers, but untrained community observations are unable to accurately pinpoint the source of this difference, which is why there are so many different guesses by community members as to what collectively distinguishes G_2 speakers from G_1 speakers. Provided that each of the proposed distinctions ('standard', 'mixed', etc.) can be adequately operationalized (see Section 3.2), the empirical work undertaken in this dissertation will elucidate which of these impressions, if any, is true.

1.2. Dialect contact in context

1.2.1. Dialect contact within established communities

Dialect contact can occur within communities that are already well established linguistically, for example when individuals who relocate to a new community speak a dialect different from the one that is locally predominant within their new community (a situation sometimes studied under the terms SECOND-DIALECT ACQUISITION [Chambers 1992; Siegel 2010] or HERITAGE linguistics [Benmamoun, Montrul, and Polinsky 2013]). While exposure to different usage models has been found to have little effect on G_1 speakers – who are generally limited in the amount of vernacular reorganisation they can perform, regardless of the degree to which they interact with the new community (Labov 2001, 415–17) – many studies have demonstrated a clear effect of such exposure on the speech patterns of G_2 individuals, who seem to have vastly more flexibility in this respect, and for whom conformity to the local norms might be more desirable (perhaps for social reasons of 'fitting in') than it is for G_1 speakers.

For instance, Payne's (1980) seminal study of the linguistic differences between children born in Philadelphia (King of Prussia) and those from out-of-state found that the latter group quickly acquired the local Philadelphia vowels, and that the density of their social network in the

new locale (measured according to the number of times a speaker was mentioned by their peers in the same study) correlated with their adoption of Philadelphia speech norms. Her results further suggest that there is a critical threshold of participation in the community above which newcomers are far more likely to adopt the local norms. Payne's study thus illustrates one possible outcome of a dialect contact situation: children reject the parental norms to instead conform to the locally predominant norms, but they do so especially well if they are closely connected to a network of local individuals. Note that local Philadelphia children do not conform to the incoming out-of-state children's norms, which suggests that accommodation in these situations proceeds unidirectionally.

Other studies of similar situations find similar effects – both in terms of direction of accommodation and in terms of the importance of connectedness to the local community. For example, there have been many studies that show if and how speakers of different dialects have adapted to each other's dissimilar linguistic patterns once these are brought into contact in a locale where another dialect is already well-established (e.g. Canadians in Southern England [Chambers 2003], in York [Tagliamonte and Molfenter 2007], and in Alabama [Munro, Derwing, and Flege 1999]; Americans in Australia [Foreman 2000]; and rural Norwegians in the urban center of Bergen, Norway [Trudgill 1986; 1994], to name but a few). Taken together, the general conclusion from these studies is that when a newcomer settles in a speech community where their dialect is unrepresented, they are able to conform to the local speech patterns – somewhat less completely if they are adults, but very proficiently if they are children – such that they ultimately mirror the existing local patterns by the time they reach adulthood (if not much sooner).

Though many of these studies are only able to examine a handful of speakers (a frequent limitation in this body of work, since locating large samples of newcomers and/or their children tends to be rather challenging), each of them focuses on newcomers integrating themselves into well-established communities where speaker interaction with other in-group members is regular and frequent. While this is not a lacuna *per se*, it does raise the question of whether the recurring conclusion of G_2 cohesion across many studies is because the communities under study in this body of work are all very similar in terms of interaction contexts and frequency. Few studies have examined the outcomes of dialect contact when the occasions for contact are both limited in terms of raw frequency, as well as contextually restricted to only one or two domains. Assessing this will directly shed light on whether degree or contexts of contact can ultimately affect the linguistic outcomes of this contact at G_2.

For example, previous research suggests that speakers must engage in regular and sustained interaction with each other in order to develop shared norms[7]. In Giles' view, individuals in the course of regular social interaction make subconscious adjustments ('accommodations') to their linguistic, paralinguistic, or nonverbal behaviour in order to increase or decrease social distance with their interlocutor (Giles and Ogay 2007, 295). Whereas Giles' initial conception of accommodation does not appear to require that interactions take place face-to-face (indeed, Giles and Ogay [2007, 300] explicitly mention the media alongside face-to-face interaction as a possible contributor to convergence), in many linguistics studies, the in-person nature of these interactions is often either explicitly emphasized or implicitly assumed to be a crucial trigger of

[7] Speakers must interact to develop shared norms, but shared norms do not *always* develop between speakers of two different dialects or languages as a consequence of interaction (Poplack and Levey 2010).

convergence (Auer, Hinskens, and Kerswill 2008; Chambers 2003; Labov 1972, 324; Samuels 1975, 90; Trudgill 1986, 54; Trudgill 2004, 158). Trudgill (1999, 6) puts it this way:

> "Face-to-face interaction is necessary before diffusion takes place, precisely because it is only during face-to-face interaction that accommodation occurs. In other words, the electronic media are not very instrumental in the diffusion of linguistic innovations, in spite of widespread popular notions to the contrary. The point about the TV set is that people, however much they watch and listen to it, do not talk to it (and even if they do, it cannot hear them), with the result that no accommodation takes place." (Trudgill 1999, 6)

The overarching objective of this dissertation is to determine whether accommodation is also possible when face-to-face interaction is available, but highly restricted (cf. Foulkes and Docherty 1999) – for example, in a community like Victoria where no 'local' dialect is established, where interactions take place between a small number of individuals, within very small social networks, or within a limited range of contextual domains (e.g. at home). Generally speaking, where opportunities for social interaction are very limited, such that social groups outside the home develop more slowly or not at all (cf. Le Page & Tabouret-Keller, [1985]; see also Kerswill & Trudgill, [2005, 200]), speakers may be more likely to retain the characteristics of the variety spoken in the home, rather than converge on the norms used by the broader community.

The Victoria francophone community provides an invaluable opportunity to assess these questions directly. If Victoria G_2 speakers do, in fact, share a way of speaking, this would

constitute evidence that even the least auspicious conditions can still lead to convergence at G_2. This, in turn, provides an opportunity to question the mechanisms that are thought to underlie convergence. For instance, Trudgill has been a strong proponent of the effects of frequency on contact outcomes in the realm of dialect contact scenarios (Trudgill 1986; 2004; 2008)[8]. In contrast, other scholars have expressed the view that social factors, such as identity or attitudes, take precedence over frequency or – at the very least – operate alongside it as equally crucial factors. Holmes and Kerswill (2008, 274) stress that "the *range of social reasons* that people adopt one form over another" (emphasis in original) is of great importance to the study of dialect contact, and should not be set aside in favour of an entirely frequency-based explanation.

In their study of Caribbean and Mainland Spanish speakers in New York City, Otheguy and Zentella (2012, 114–15) provide a concrete example of this when they note that certain "numerically superior" lexical variants do not dominate in converging dialects, precisely because they are socially evaluated as stigmatised. Similarly, Thelander's (1983) study of migrant groups in Sweden found that speakers tended to avoid dialect convergence when their original dialectal form corresponded to the standard form, instead of converging on the more frequent, but nonstandard, form. Dodsworth (2017, 332) summarises several studies where "the 'winning' variant is more frequent in the initial population of speakers" (supporting the 'frequency' stance: [Mufwene 2001; Thelander 1979; Trudgill 1986; Trudgill et al. 2000]) but also several others where "the 'losing' variant is marked in some way" (supporting the 'social evaluation' stance: [Kerswill and Williams 2000; Moag 1977; Trudgill 1986]). My study is designed to account for the possible tension between frequency and social evaluations, by incorporating cohorts of

[8] A view which has clear associations with Labov's PRINCIPLE OF DENSITY (see Trudgill [2008, 252]).

speakers from dialect regions that are not equally represented within the community as a whole on the one hand, as well as a number of extralinguistic (e.g. level of education) and sociostylistic (e.g. speech style) factors on the other (see Section 3.4).

1.2.2. Dialect contact in new territory

The evidence from studies of established communities, summarised above, would suggest that G_2 speakers generally conform to the established local dialect, even as G_1 speakers maintain the distinct usage patterns that characterise their original dialect(s). But – as far as we know – no such established dialect exists for Victoria. French-speakers of various different dialects frequently move in and out of the region (as described more fully in Chapter 2), and this mobility causes impermanence within the community and a relatively unstable linguistic situation. In other words, new francophone arrivals to Victoria find themselves in essentially 'unsettled' linguistic territory (at least as far as French is concerned; English being well established as the majority language in the region), in the sense that there is no enduring variety of French that is entrenched in the Victoria region. Though somewhat less common than the dialect contact scenarios outlined in the preceding section, there are still several studies of situations where speakers of different dialects relocate to the same new region, previously unsettled by speakers of the language in question.

The outcomes of these situations seem to differ to an extent from those described in Section 1.2.1. In new territory, speakers experiencing dialect contact may opt to incorporate elements of several models – perhaps giving new sociolinguistic functions to existing forms, alternating between forms that exist in two competing models, or 'mixing' elements from multiple models to create new, 'hybridized' forms (cf. Trudgill's [1986; 2004] NEW-DIALECT FORMATION or Kerswill's [2013] KOINEISATION). For example, the linguistic *méli-mélo* ('hodgepodge')

attributed to G_2 speakers in (1), might give the impression that we are dealing with something like 'interdialect' or 'intermediate' forms. These consist of phonetic variants that are acoustically intermediate between the variants associated with each of the dialects in contact (also known as 'fudged' variants [Chambers and Trudgill 1998), or of additional variants not present in either of the precontact varieties (Trudgill 1986, 60). Alternatively, the *méli-mélo* could correspond to what is known in the dialect contact literature as a 'koine' (cf. Kerswill 2013; Kerswill and Williams 2000; Siegel 1985). To put it concisely, koines may arise when G_1 speakers from different dialect regions migrate to a territory where neither of their dialects is established, then produce a second generation of speakers whose speech reflects a 'mixture' of the G_1 dialects or (sometimes) incorporates innovations. The initial G_1 mixture is additionally said to undergo SIMPLIFICATION (which Kerswill and Williams [2000, 67] define as "a loss of irregularity in morphology, a reduction in the number of grammatical categories, and an increase in invariable word forms") and FOCUSING (which Le Page and Tabouret-Keller [1985] equate with an increase in regularity). In addition to the basic requirement that speakers interact frequently, Kerswill and Williams (2000, 75) propose several more factors that affect the likelihood of koineisation. To paraphrase them here, they suggest that a higher ratio of adults to children in a given dialect contact scenario will promote more rapid simplification (which is, in their view, only an initial stage of koineisation); conversely, more children in the mix will inhibit simplification. A "highly normative approach to mother-tongue literacy" is also thought to inhibit simplification (*ibid.*, 75). Factors that promote focusing include demographic features that act to enhance socialization or the formation of new social networks between children; factors that inhibit focusing include the "degree of linguistic difference" and "complexity" of features involved in the mix (*ibid.*, 75).

As is perhaps evident from my liberal use of direct quotes in the above description, I hesitate to wholeheartedly espouse this framework and its associated terminology for several reasons. Firstly, the theorization that underlies this perspective largely seems to assume that forms can be *categorically* attributed to one variety or the other, which is very difficult to reconcile with the inherent variability of natural speech at different levels of the grammar. Notions such as 'interdialect forms' or 'mixing' can really only be invoked if the analyst can be *certain* of the categoricity of the phenomenon at hand – something which is not always convincingly demonstrated. Secondly, there is generally little objective evidence provided to support the use of terminology such as 'complexity' or, conversely, 'simplification' – these epithets are used freely, but tend to be unqualified and vague. For instance, one definition of 'simplification' used by this framework (there are several) – "a reduction in the number of variable categories" (Kerswill and Williams 2000, 67) – is effectively indistinguishable from many types of garden-variety internal change that are *not* specifically triggered by dialect contact or multiple-model environments. Furthermore, a reduction in *categories* might yield no differences to the underlying *conditioning* – if a speaker uses a variant less frequently but retains the full spectrum of underlying variability for the other variants in a given system, it seems specious to conclude that this is 'simplification'. As long as variant selection is governed by implicit conditioning, and especially if these conditions are shared across the community in question, it is inappropriate to talk of simplification.

Finally, in order to avoid equating *variability* with *change* (a pitfall to be avoided in contact research more generally; see Poplack and Levey [2010]), it is imperative to identify and adduce benchmarks of closely-matched *non*contact data to serve as points of comparison with the post-contact variety (e.g. as described in Section 3.2). Because studies of dialect contact in new

territory are often unable to obtain data on anterior stages of the dialects they are studying – for the simple reason that such data may not exist – any conclusions of 'change' must be viewed with a critical eye. Indeed, this is the main thrust of Poplack and Levey's (2010) 'cautionary tale' against hasty conclusions of contact-induced change: the simple fact that two languages or dialects may have been in contact at some point or are currently in contact is not sufficient to conclude that similarities between these languages or dialects occurred *as a result of* contact. The onus is on the analyst to test whether such similarities could have arisen through parallel internal developments (which Poplack and Levey term 'drift'; *ibid*., p. 397), by first assessing the underlying structural patterns for the suspected change within the source and recipient varieties, then by proving that the change was not already present in a pre-contact stage of the recipient variety but was present in the source variety, and finally by ruling out recipient-internal explanations for the change (*ibid*., p. 410). Applying this method to various candidates for contact-induced change has shown that such changes are "*not* an inevitable, nor possibly even a common, outcome of language contact" (*ibid*., p. 412). These principles should be transposed onto the study of *dialect* contact as well, to the extent that it is possible.

These methodological issues do not necessarily diminish the findings of studies that examine dialect contact in new territory, since some of the findings can be described without the uncritical use of terms like 'simplification' or 'mixing'. For instance, Trudgill (2004, 101) observes that G_2 New Zealand speech is "unlike [ ] stable situations where children normally acquire the dialect of their peers" in that "the role of adults [as linguistic models] will be more significant than is usually the case." He goes on to describe the speech of certain individuals in some detail, drawing phonetic and phonological parallels between their speech and that of their parents, and differences between their speech and that of other, generationally-matched, individuals in the

community in question. Within the context of the Victoria situation, this finding would suggest that the absence of an existing stable variety to acquire might cause G_2 speakers to talk more like their parents than their peers. The retention of parental norms would also cause a certain degree of variability between individuals, as attested by Trudgill ([2004, 106]: "people who have grown up in the same place at the same time may differ very markedly from one another"). That is, the community undergoing dialect contact in new territory might not be as immediately susceptible to convergence as communities where there is an established dialect to converge on. These are important considerations for the design of the present study, which is why I assess the speech of G_1 speakers alongside that of G_2 speakers in order to determine whether there are, in fact, similarities or divergences between them.

Trudgill (2004, 113) also argues that, in later stages of new-dialect formation (one possible outcome of dialect contact in new territory), "large numbers of variants from the different dialects involved in the mixture are reduced in number, until usually only one variant remains for each variable." At face value, this would seem to be a rather extreme and misguided stance, since it effectively suggests that dialects arising through convergence in new territory are completely invariant. But within the context of Trudgill's research, this is suggesting that the multitude of available allophones for a given phoneme in the early days of the New Zealand settlement eventually consolidated into a single shared phonetic system. Although this phonetic consolidation takes place rather quickly, all things considered, it definitely takes longer than similar processes of convergence in established communities: where Payne's King of Prussia children acquired a new Philadelphia vowel system within a few years of their arrival, Trudgill's New Zealanders take three generations to achieve a similar stability. Since the Victoria situation is characterised by extreme instability, as well as a lack of any established variety to serve as a

model for newcomers, it might be the case that G_2 development of shared phonetics is more or less severely hindered, such that they have not achieved a cohesive system like the one attributed to third-generation New Zealanders.

1.2.3. Dialect contact within the Victoria francophone community

In terms of demographics and social structure, there are many clear parallels between the communities described in my review of the dialect contact literature and the francophone community in Victoria. The purpose of the present section is to isolate and describe the models in the linguistic environment of Victoria G_2 speakers that might be most likely to influence their speech. Observations provided by the participants in (1)–(4) seem to suggest that G_2 speech is characterized by a 'mix' of French from Quebec or Ontario (Laurentian French, 'LF') and French from France (Hexagonal French, 'HF')[9]. These two varieties are known to differ in terms of both their structural base, their internal conditioning, and their patterns of sociostylistic stratification, in ways that will be unpacked further in Section 3.3. They therefore provide two *competing* usage models to G_2 speakers. The possibility of influence from one or both of these sources is certainly supported by the demographics of the community: around 73% of the population comes from a Laurentian French-speaking region outside of BC (Quebec or Ontario), and another 15%

[9] The terms 'Laurentian French' and 'Hexagonal French' are used throughout this study as a shorthand to designate speakers that share certain identifiable phonological and morphosyntactic characteristics which are consistently associated with the broader Laurentian or Hexagonal dialect regions. The terms are not meant to imply that every LF or HF speaker shares every distinguishing characteristic, but the relative internal homogeneity within each these two regions (at least as far as the characteristics of interest in this study) validates the consolidating labels here.

comes from outside Canada – mostly France (cf. Chavez and Bouchard-Coulombe 2011). In fact, only around 12% of the province's francophone population was born in BC.

The fact that the dialect contact in Victoria happens to occur between LF- and HF-origin speakers is fortuituous in a way, because it allows for a more direct examination of whether the factor of frequency or that of social motivations is a bigger contributor to G_2 convergence. Per Trudgill's strong frequency stance, forms occurring at "too low a level of frequency to be noticed" (Trudgill 2004, 111) will not ultimately be retained in G_2 patterns. This view is echoed in a great number of studies that examine the effects of contact between speakers from different dialect regions (e.g. Chambers and Trudgill [1998]; Kerswill and Williams [2000]; Payne [1980]; Trudgill [1986; 2008] – to name but a few), which hold that the direction of convergence (if this outcome can even be demonstrated to have occurred) is entirely predictable from the relative proportions of each dialect group in contact. In Victoria, LF-origin speakers outnumber HF-origin speakers at a ratio of around six to one, which would make the LF model a decidedly more likely one for speakers to adopt according to this view.

However, the choice between HF and LF also has great social significance in the Canadian context. The HF variety is imbued with OVERT PRESTIGE, which refers to varieties that are "recommended by powerful social institutions" (Crystal 2003, 115) or "promulgated by the agents of standardization in society" (Guy 1988, 51). Historically, overt prestige in French has been centred on the variety spoken in the *Île-de-France* region (Kircher 2009), but recent standardisation efforts in France have effectively broadened the association of prestige to the entire country, as metropolitan Parisian French has spread throughout formerly linguistically distinct regions (see Armstrong [2001, 5]). Even though spoken HF certainly does not align one-to-one with the rules of grammatical prescription (see Section 3.3), it is often perceived as an

instantiation of 'standard' French nonetheless. For instance, speakers of HF have been historically rated by LF-origin individuals as evoking higher social status, intelligence, and ambition (cf. d'Anglejan and Tucker 1973), and echoes of these evaluations persist even in more contemporary works (Armstrong and Pooley 2010; Dumas 2001; Francard, Lambert, and Berdal-Masuy 1993; Oakes and Warren 2007; Reinke 2005; etc.).

Studies that examine the evaluations given by francophones to different guises (e.g. 'European' French versus 'Quebec' French, as in Genesee & Holobow [1989] and Fuga [2002]) find that European French guises are rated higher in status (i.e., overt prestige) than Quebec French guises by anglophones and francophones alike – a finding also replicated in Maurais (2008) and Kircher (2009). Bouchard and Maurais (1999) find that francophones express fewer positive attitudes towards Quebec French than anglophones do, and that participants' linguistic preference is generally for *Radio-Canada* French (i.e., 'standard' Laurentian French). Kircher (2009, 106) points out that "during the first year of the 21st century, francophone Quebecers continued to evaluate QF [Quebec French] more negatively than EF [European French] with regard to status" – suggesting that, although the last forty years or so have certainly seen a significant improvement in the evaluations of LF, the divide between HF and LF persists to some extent.

Within the Victoria context specifically, many participants express attitudes that would indicate a somewhat negative evaluation of LF as it is spoken in Quebec (6)–(8). This might at least partly be due to the fact that many G_2 speakers seem to have had negative experiences with Quebecois people who doubt their francophone identity (9), but could also be explained by the evaluative imbalances between HF and for LF that were described above.

(6) Moi j'ai amélioré beaucoup mon français en sortant du Québec. [VFC.LF$_1$.Albert.01/53/51]

'I really improved my French by leaving Quebec.'

(7) Comme eux ils ont beaucoup d'anglicismes, mais– mais je pense, comme, nous on a tellement fait un effort, surtout à l'école de, comme, juste parler en français que des fois je pense qu'on a moins d'anglicismes. [VFC.LF$_2$.Héloïse.01/18/21]

'Like they have lots of anglicisms, but– but I think, like, we made such an effort, especially at school to, like, just speak French that sometimes I think we have fewer anglicisms.'

(8) Je dirais que notre vocabulaire– notre français est mieux que celle au Québec parce qu'on a moins de– de, comme, slang ou on utilise vraiment les mots appropriés quand– tandis qu'eux autres ils vont juste sortir plein d'affaires. [VFC.LF$_2$.Geneviève.01/34/29]

'I would say that our vocabulary– our French is better than the one in Quebec because we have less, like, slang or we really use the appropriate words when– whereas they just come out with all kinds of things.'

(9) Les québécois nous disaient uh– moi ils me posaient « tu viens d'où toi? T'es pas québé– » « Ah, okay, okay, » je l'ai convaincu finalement mais fallait argumenter avec lui parce mon accent, il dit « bien, tu-sais, un peu français, un peu anglais. » [VFC.HF$_2$.Frédéric.01.00.03]

'The Quebecois would say to us uh– they would ask me "where are you from? You're not Quebe–" "Ah, okay, okay," I finally convinced him but I had to argue with him because [of] my accent, he says "well, you know, a bit French, a bit English."'

While the results from Kircher (2009), as well as the other studies on attitudes towards HF and LF, suggest that HF might have more overt prestige or evoke higher status than LF, LF unquestionably has more COVERT PRESTIGE, as it ranks above HF on dimensions of solidarity. Varieties with covert prestige are not typically viewed as 'standard', but are nonetheless "positively valued, emphasizing group solidarity and local identity" (Crystal 2003, 115). On the one hand, choosing a model with *overt* prestige, such as HF, might be perceived as distant, snobbish, or condescending, whereas opting for an LF model signals group solidarity, and is therefore quite possibly a more appealing option to G_2 speakers – who, after all, belong to the broader French-Canadian context. On the other hand, the HF option (or at least, something perceived as more 'standard' than garden-variety, colloquial LF[10]) might have greater appeal because of the connection between prestige and ethnolinguistic vitality – a major concern for francophones in minority contexts. Personal communication with many Victoria francophone individuals[11] (especially G_1 individuals) reveals the pervasive but erroneous belief that the 'quality' of spoken French is what ensures its vitality. It is true that varieties of French with high

[10] The speech of *Radio-Canada* announcers, as an instantiation of what is sometimes called 'Standard Quebec French', might fall into this category (Bourhis and Lepicq 1993; Bouchard and Maurais 1999, 96; Kircher 2009; Rochette and Bédard 1984). Unlike spontaneous spoken HF, *Radio-Canada* French does approximate the grammatical standard quite closely in some ways, including for some of variables assessed in this study (e.g. *avoir/être*, described in Chapter 5). Studies of Radio-Canada speech (e.g. Bigot 2008; Villeneuve 2017) demonstrate that this variety uses a higher rate of 'standard' features, compared to everyday, informal LF speech.

[11] Since 2018, I have given interactive workshops and lectures on linguistic insecurity to parents, teachers, students, and community associations throughout the province of British Columbia, which has provided me with an insider's view of the beliefs and opinions held by many different sectors of the population. Though this is anecdotal evidence, it has shaped my understanding of the local francophone community and has guided the research process undertaken here.

ethnolinguistic vitality (i.e., where French is the exclusive or vastly predominant language of communication) also tend to have higher prestige (e.g., Giles and Johnson 1987, 71; Giles and Ogay 2007, 299) – HF has high prestige as well as high ethnolinguistic vitality, for example – but it does not appear to be the case that prestige directly *triggers* ethnolinguistic vitality in any way. Nonetheless, the association of prestige with higher vitality (or the belief that 'good' French ensures its survival) might lead Victoria francophones to gravitate towards a higher-status variety – potentially even as a conscious strategy to increase francophone vitality.

While it seems clear that HF and LF do have different evaluations of prestige or status, it is less clear whether prestige can stand alone as an explanation of linguistic behaviour. For one thing, sociostylistic evaluations cannot be assumed, a priori, to be uniform across all sectors of a community: speakers of HF and speakers of LF in Victoria might very well differ in their attribution of 'prestige' to some variables and 'stigma' to others[12]. After all, the sociostylistic evaluations attributed to different variables, or to the linguistic contexts in which a particular variant might be used, are not fixed properties of those variables or variants. As Milroy (2012, 572) puts it: "whether a linguistic form is believed to have high or low prestige depends, not on its linguistic shape, but on the perceived social status or importance of the speakers who use it." This suggests that the social context of the Victoria francophone community might yield different sociostylistic evaluations for certain forms than within the contexts of France or Quebec. Thus, using the concept of 'prestige' as an explanation for linguistic behaviour is not

[12] It should be noted, however, that shared sociostylistic evaluations form a key component of the Labovian definition of the 'speech community'. Per this definition, differences in sociostylistic evaluations would mean that HF and LF form separate speech communities (or at least, do not form one unified speech community).

necessarily problematic in and of itself, but any prestige-related factors (e.g. those operationalising speech style, social class of the speakers, level of formal education, etc.) must be convincingly established for the sample before drawing conclusions as to their effects. Doing so will shed some light on the question of whether attitudinal factors may have played a role in shaping G_2 speech and will therefore contribute indirectly to the question of whether attitudes or frequency are more important in convergence.

In addition to the two implicit usage models provided by HF and LF, a third model is provided to G_2 speakers within the classrooms of the francophone school, where grammatical rules are explicitly taught as part of the curriculum starting in around grade three (roughly, ages 7-9). The prescriptive rules that are taught from textbooks or modeled by teachers during lessons[13] often directly contradict the implicit usage norms provided by HF and LF speakers alike:

> "Authors [of pedagogical texts] still seem to be reluctant to put 'incorrect' French into written form ... students might not read in the textbooks what they hear, might be told what they should hear and then see half of it reproduced, or might even be told right away that they will not be taught what they are asked to listen to, resulting in an altogether complex, schizophrenic representation of the 'norm'." (Etienne and Sax 2009, 598)

[13] Previous studies (e.g. Mougeon and Rehner 2019; Poplack 2015) show that the French spoken by teachers in the classroom often incorporates more 'standard' features, uses them at higher rates, and/or avoids more vernacular features, especially when those features are subject to strong prescriptive stigma.

However, prescriptive rules do have certain advantages which might make them especially accessible to G_2 speakers in this situation. They are exceptionally stable: grammatical prescription is slow to change and therefore relatively uniform across many generations, and the same rules apply regardless of the dialect of French one speaks (i.e., a speaker of HF is taught from the same grammar books as a speaker of LF). Prescriptive rules are also arguably very salient, since they are explicitly taught and rehearsed at school over a number of years. Perhaps most important, however, is the fact that, for a large sector of the G_2 population, the classroom represents one of the only contexts in which French is regularly used. Thus, 'school' French (school-F)[14] might hold more sway over HF or LF usage norms, since the latter might get less airtime in the daily lives of many G_2 speakers.

1.3. Research questions and theoretical framework

Given what has just been described – the competing dialect models in the Victoria environment, the lack of community stability and face-to-face interaction, the sense of cultural solidarity and identity that is encouraged by the school – the Victoria situation invites investigation along several lines. Has the socio-cultural cohesion of G_2 speakers translated into linguistic cohesion? Is it the case that G_2 speakers have instead converged on their parents' usage norms, thus perpetuating the dialect diversity that characterises G_1 speakers? If a unified G_2 variety *does* exist, does it resemble any of the G_1 varieties present in the linguistic environment – and if so, *which*? Alternatively, is there evidence that the school has played a role in shaping G_2 patterns, given its

[14] This term is used throughout to designate the abstract prescriptive rules for the use of the variables studied here, which were culled from meta-analyses of grammatical prescription (particularly Poplack et al.'s [2015] study of 163 French grammars across five centuries' worth of publication), grammar books, and pedagogical manuals for L1 and L2 French instruction (see Section 3.2).

central role in minority-language transmission and as a community hub? Confirming whether G_2 speakers have innovated new patterns relative to those that exist at G_1, and whether an institution such as the francophone school might play a key role in the development of these patterns has implications for our understanding of the conditions that contribute to shared implicit norms for speech, as well as how the social structure of a community and the context(s) in which a language is used can affect the development of these norms. The results of this study can advance our understanding of the mechanisms or contributing factors that underlie convergence in dialect contact (e.g. frequency vs. attitudes or other social motivations), but also speak to the questions of language transmission in minority contexts, linguistic discontinuity within otherwise socially-cohesive communities, and the relationship between contextual restriction and functional restriction.

My examination of these questions is couched within the theoretical and methodological apparatus of the COMPARATIVE VARIATIONIST framework (Poplack and Meechan 1998; Poplack and Tagliamonte 2001; Tagliamonte 2006; 2012). This framework enables the analyst to assess variability and change – both key elements in the situation described above – even in the absence of diachronic data, through analysis of spontaneous speech at the synchronic level. The major analytical construct of this theory is the LINGUISTIC VARIABLE (Labov 1966; 1972; 1984), which is comprised of all the linguistic forms (called VARIANTS) that have the same referential meaning within the same linguistic or structural context. This context, termed the VARIABLE CONTEXT, is carefully delimited by the analyst such that it encompasses all the linguistic environments where the variability is attested for a given variable, as well as all of the environments where the variability *could have* occurred but did not. In addition, it excludes all the environments where the variability *cannot* occur. In this manner, the full spectrum of variability in a given context is

accounted for and nothing is erroneously left out, but contexts outside the envelope of variability (i.e., those where the forms in question do not express the same referential meaning) are not considered. In order to determine what motivates the choice between the available variants, the next step is to identify factor groups that can be reasonably hypothesized to contribute to the selection of one variant over the other(s) within the variable context. In Guy's (2018, 2) words: "every independent variable [factor group] included in a quantitative analysis entails a hypothesis that it has some effect on the dependent [linguistic] variable." Discovering which elements of the linguistic and extralinguistic context affect variant choice gives us a view onto the structure of the variability, which can also be construed as the grammar that underlies the variability. Each factor group is coded for the categories that are relevant to the hypotheses on variant selection, and the set of codes is then submitted to a statistical program for multivariate analysis specially designed to handle the eccentricities and imperfections of spontaneous speech data (*Goldvarb Yosemite: A Multivariate Analysis Application for Macintosh*; Sankoff, Tagliamonte, and Smith 2015). The distribution of variants according to the various factors proposed to influence their occurrence, as well as the statistical confirmation of each factor's (un)importance provided by the multivariate analysis, are key to understanding the underlying patterns of variation within the implicit usage norms of a given community.

The use of multivariate analysis is essential in this study, because most of the variables used to determine a) whether G_2 speakers share patterns, and b) which models have contributed to these patterns contain virtually the *same variants* in all models. In other words, a given variable might feature exactly the same variants in HF, LF, and school French, and therefore it is impossible to rely on the simple presence or absence of a given variant in G_2 speech to determine which of these three models has contributed it. This kind of situation requires a method which is

capable of handling variability, but also one which accesses something more than the categorical presence or absence of each variant. The comparative variationist method scores on both counts. Where possible, I also privilege underlying patterns (termed 'conditioning') over rates. The reliance on rates alone to compare two different studies or varieties has been problematized in much of Poplack's work (with Dion [2009]; Lealess & Dion [2013]; Levey [2010]; Tagliamonte [2001]; and more), since rates of any given variant may fluctuate for entirely non-linguistic reasons (Poplack and Tagliamonte 2001, 92). Careful handling of variables in this manner is especially important in a community where fluctuations in dialect mixture with each loss-and-renewal cycle are likely to shift the rates of a given feature – perhaps obscuring the fact that the underlying conditioning has not changed.

I identified four diagnostic variables that will clearly illustrate differences between HF usage, LF usage, and school French. Each consists of an alternation between two or more variants – between lenis and fortis realizations of coronal plosives in assibilation, between the auxiliaries *avoir* and *être* in compound tenses of the 'Vandertramp' subset of verbs, between periphrastic and synthetic future constructions in future temporal reference, and between subject-verb inversion, rising intonation, the question particle *-tu*, and the question phrase *est-ce que* in yes/no questions – each of which will be discussed in more detail in Section 3.3. Note that these variables are not always perfectly diagnostic of a *three*-way difference between LF usage, HF usage, and school French, but each of them tests at least one axis of comparison between these competing models (see Table 1.1). For example, there are no prescriptive rules or normative associations for assibilation – the assibilated pronunciation that characterises LF is simply recognized as the customary option for LF; it is not *pro*scribed in modern grammar books, nor even mentioned in contemporary pedagogical literature, and is used in both the most informal

and the most elevated or formal speech contexts in Canada (e.g. by *Radio-Canada* announcers in televised interviews). As such, assibilation can only diagnose whether G_2 speakers align with the HF usage norm or the LF usage norm. It cannot diagnose whether G_2 speakers align with school French.

Table 1.1. All diagnostics used to evaluate G_2 speech

	HF vs. LF	HF/LF vs. SCHOOL-F
Assibilation	Rates only	N/A
Avoir/*être*	N/A	Rates / conditioning
Future temporal reference	Rates / conditioning	Rates / conditioning
Yes/no questions	Rates / conditioning	Rates / conditioning

This ensemble of variables allows me to test different levels of linguistic structure, thereby allowing me to build a more revealing picture of the influence that various input sources may have had on G_2 speech, to draw conclusions about whether there is such an entity as 'G_2' speech, and if so, to determine which models G_2 speakers have converged on. Phonetic/phonological variables, with their relative categoricity and salience, might lead the analyst to very different conclusions than if they had also studied morphosyntactic variation (cf. the question of 'simplification' mentioned earlier). Adducing multiple variables in this manner will provide a far more complete picture of convergence (including its complex patterns of variability) than if only one type of variable had been selected. Assessment of multiple variables is especially important because the unique sociolinguistic context of the Victoria francophone community may yield any number of different outcomes for G_2 speakers. For instance, G_2 speakers may preserve dialect diversity of G_1 speakers by aligning themselves with the implicit usage norms that are/were

present in their home. In other words, a G_2 speaker whose parents are speakers of LF might employ only LF usage norms, whereas speakers with HF-origin parents might employ only HF usage norms. On the one hand, this would be a surprising outcome – both in consideration of the community impression that G_2 speech does *not* align with G_1 speech, but also given previous findings in the broader field of language transmission which suggest that G_2 speakers conform to their peers, not their parents. On the other hand, the isolation and instability of Victoria's francophone community, coupled with the fact that many G_2 peer groups communicate in English, might militate against the development of any unified/shared G_2 norms (cf. Trudgill's [2004] findings for G_2 New Zealanders).

Alternatively, G_2 speakers may instead converge on shared implicit usage norms that a) are not distinguishable from one of the G_1 options – that is, G_2 speech might hypothetically align with the implicit usage norms of either LF or HF (i.e., the two main sources of usage norms that are present in their linguistic environment) – or b) correspond more closely to school French than to either of the usage norms (perhaps by mirroring rates associated with 'standard' French for each of the variables). The former outcome strongly resembles the vernacular re-organization described by Labov (2001, 416–17), whereby newcomer children acquire a local dialect "almost completely" within a short time of their arrival – provided that there is an established, stable local dialect to acquire. I argue that population mobility and turnover have prevented the formation of such an established French dialect in Victoria. However, the G_1 individuals in Victoria are *largely* from LF-speaking regions. It might be the case, then, that the LF population – though its individual speakers are ephemeral – nonetheless acts as the 'local' (or possibly 'supralocal') variety. Then again, since G_2 French usage is largely confined to the school context, their speech might be coloured by the prescriptive rules of school French or influenced by

associations of a particular variant with higher social status or greater education. It might even be easier for G_2 speakers to converge on school French because, although the standard is certainly not immutable nor perfectly consistent (cf. Poplack et al. 2015), it certainly changes far more slowly than speech, thus making it a potentially more stable choice for G_2 speakers to 'latch onto' in the absence of the same stability in the usage norms to which they are exposed.

The social circumstances that distinguish G_1 and G_2 individuals, alluded to above and elaborated in more detail in Chapter 2, lead me to hypothesize that G_1 individuals will retain their original implicit usage norms, thus maintaining dialect diversity overall at the G_1 level. In contrast, given the widespread impression that G_2 individuals speak in an identifiable way, I propose that they will not perpetuate this dialect diversity into a second generation and instead will converge on shared usage. Given the importance of the school as a primary context of French transmission and maintenance, as well as the relative stability and uniformity of school French (compared to usage norms that might fluctuate along with the mobility of community members), I further propose that G_2 speakers' shared norms will reflect a greater influence of school French than usage. In other words, I believe that the circumstances in which G_2 individuals are exposed to and use French will ultimately affect their linguistic output, such that G_2 speech reflects increased use of prescriptively-mandated or high-status forms.

1.4. Roadmap

This dissertation continues in Chapter 2 (*The Victoria francophone community*) with a detailed description of the community characteristics that are likely to affect linguistic outcomes, including those alluded to throughout this chapter (isolation, instability, etc.). Understanding the social context in which French is used in Victoria will prove to be illuminating when it comes to unpacking the results. Chapter 3 (*Data and method*) opens with a detailed description of the

Victoria French Corpus, which I constructed for the express purpose of this study, and continues with a more comprehensive overview of the methodological machinery I employ. Chapter 4 (*Assibilation as a window onto competing usage norms*) explores the conflict between HF and LF usage that G_2 speakers must resolve. Chapter 5 (*Avoir/être: A salient conflict between usage and the standard*) tests a different angle of the same question: whether subconscious implicit usage or salient, explicitly-taught grammatical rules of French win out in G_2 speech. Chapters 6 and 7 both treat three-way conflicts between LF usage, HF usage, and school French, but the conflict in the former (*Future temporal reference: A three-way conflict with non-salient variability*) proceeds completely below the level of conscious awareness, whereas the conflict in the latter (*Yes/no questions: A three-way conflict with salient variability*) includes an element of salience that might allow speakers to more readily perceive and adopt this complex system. Each of Chapters 4 to 7 constitutes a full quantitative analysis of one of the four diagnostic variables selected for this study (including a literature review, the variable and its variants, the circumscription of the variable context, the results, and an interim discussion). The findings from these chapters are then amalgamated and further discussed in Chapter 8 (*Finding the big picture: Summary and general discussion*), which aims to clearly outline the 'big picture' gleaned from the combined analysis of so many different diagnostics, as well as to connect these findings to the social context of the Victoria francophone community.

Chapter 2. The Victoria francophone community

If it is true that G_2 speakers have managed to converge on cohesive usage norms, as the impressionistic assessments in (1)–(5) suggest, then this would be a remarkable achievement indeed, given the social conditions under which this acquisition would have taken place. This chapter outlines some relevant characteristics of the Victoria francophone community: the lack of establishment and permanence, the contextual restriction of French to a narrow range of social domains, the barriers to contact between speakers of different dialects that existed for the better part of the 20th century, and the importance of the school to the maintenance and vitality of G_2 French.

2.1. Establishment and permanence

The Victoria francophone community has actually been 'established' since at least the mid-nineteenth century (Agnew et al. 1987), but this stability over some 200 years is illusory. As generations of residents eventually succumb to assimilation (almost certainly by the third generation, if not sooner; cf. Chavez and Bouchard-Coulombe [2011]), new arrivals from outside the region continually pour into the community, thereby 'renewing' it in some sense. In other words, though the francophone community in Victoria appears to have a long history, their linguistic roots continue to run very shallow. Many G_1 arrivals to Victoria do not settle there for very long – only rarely do they stay more than a few years (Chavez and Bouchard-Coulombe 2011) – and if they do stay, their scattered distribution across the region (described in (10) as *îlots* 'islets' by one community member) seems likely to inhibit the regular social interaction that would be necessary for accommodation.

(10) Alors uh, moi– la façon dont je– je percevais la– la francophonie à Victoria, c'était des îlots […] donc ça serait bien, tu vois, si à Victoria on avait un quartier francophone. [VFC.HF$_1$.Fernand.00.49.17]

'So uh, me– the way I– I perceived the– francophonie in Victoria, it was islets […] so it would be nice, you see, if in Victoria we had a francophone neighbourhood.'

Where previous research shows that prolonged contact between dialects tends to (eventually) push speakers to converge on shared norms, it does not initially seem as though there is sufficient permanence in Victoria's francophone community for this to have occurred. The situation in Victoria is thus in stark contrast to many previously-studied situations of dialect contact, where the first generation of arrivals settles in the new region more or less permanently, and where subsequent generations acquire language in this more stable context. In Victoria, G$_2$ speakers are acquiring and using French as the demolinguistic landscape continually shifts around them.

In addition to the rapid population turnover, the community is subject to slow population growth. Net francophone migration to BC is very low: between 2001 and 2006, around 7,400 francophones arrived in BC from other provinces, but 7,000 francophones left BC within the same time period – yielding a net migration of only approximately 400 francophone individuals over *five years*. Compounding the slow growth of the population is the fact that the community is small and exists within a strong English majority, which acts to limit opportunities for speaker interaction in French, but also depresses the rate of transmission of French to children (around 26%, cf. Chavez & Bouchard-Coulombe [2011, 16]) and reduces the proportion of families for

whom French is the primary home language (only 27%, cf. Table 3.4.6; Chavez and Bouchard-Coulombe 2011, 26). In other words, francophones (both G_1 and G_2) who have long been established in Victoria are far more likely than new arrivals to be undergoing some kind of attrition, but since immigration is more or less constant, the influx of new arrivals acts to replace the losses sustained by this attrition, rather than to reinforce the numbers of the existing population. The population thereby maintains a kind of stasis where the number of francophone individuals remains roughly the same, but changes in progress elsewhere in the francophone world are continually being re-introduced via the implicit usage norms of new arrivals.

2.2. Francophone domains

Victoria is a predominantly English-speaking city, and the ubiquitous presence of English in the public sphere contributes to the general invisibility of the francophone population. Chavez and Bouchard-Coulombe (2011, 27) report that 99.9% of BC francophones live in an area where francophones represent 10% or less of the population. In Victoria, francophones represent roughly 2% of the total population. Furthermore, 99% of public-sphere interactions in BC take place in English (*ibid.*). The effects of being such a profound minority might have been significantly mitigated within a geographically-bound francophone neighbourhood where community members could continue daily usage in French – as in many other minority francophone enclaves across Canada – but this is not the case in Victoria. There haven't been any predominantly francophone neighbourhoods in Victoria (or indeed, in BC) since the turn of the 20th century. The municipality of Esquimalt[15] contains the highest concentration of

[15] A sizeable proportion of the francophone population is composed of military personnel who are posted to CFB Esquimalt, Canada's Pacific Coast naval base, located immediately west of downtown Victoria.

francophones within Greater Victoria at 3.3% (Chavez and Bouchard-Coulombe 2011) – a *long* way off from being a predominantly francophone neighbourhood. For many community members, the lack of a francophone neighbourhood or other strong geographic concentration of francophones is explicitly identified as a significant barrier to the vitality of French in Victoria, since it greatly reduces the opportunities for daily interaction in French that are essential to language maintenance.

(11) Tu-sais ça c'est ce que je pourrais déplorer ici de la communauté francophone, c'est qu'on se connait pas assez. [VFC.LF$_1$.Albert.02/18/54]

'You know that's what I could deplore here of the francophone community, it's that we don't know each other enough.'

(12) On est beaucoup éparpillé. On a toujours été éparpillé. [VFC.LF$_1$.Denise.01/34/32]

'We're very scattered. We've always been scattered.'

(13) Ça mixe pas comme– uh– tu voudrais. [VFC.LF$_1$.Bob.01/35/46]

'We don't mix like– uh– you would like.'

Around 10% (≈440/4400) of CFB Esquimalt's personnel is francophone (Military Family Resource Centre website, 2020), which is equivalent to close to 30% (≈440/1600) of Victoria's total francophone population. The military has also historically played an important role in French-language education in Victoria: Brodeur was originally established as a school for the children of francophone military personnel.

Furthermore, as there are extremely few public establishments where French is the predominant language in Victoria[16], French will likely not even be *understood* by staff or patrons of most businesses – never mind spoken as the primary language.

French in the public domain is really only present in sporadic francophone cultural events throughout the year, or in francophone institutions such as the school or (to a much lesser extent) the church[17]. With English as the exclusive language of public expression, the chances of interaction between francophones are significantly reduced, and many established residents have necessarily created a social network from their pool of local anglophone contacts. In some cases, G_1 speakers in Victoria will maintain francophone ties with their homeland social network, but the bulk of their *local* social network is likely to be anglophone, not francophone – simply by virtue of sheer predominance of the former. As such, for the French-speaking people in Victoria who do *not* partake in the cultural events or visit the institutions (including those who may have arrived in the region in the years before these institutions existed or those whose children attended the school but who have since graduated), either French will be relegated entirely to the private domain, or speakers will simply extend English to a wider number of contexts (including in the home).

[16] The francophone directory of the *Fédération des francophones de la Colombie-Britannique* lists seven professionals (including lawyers, accountants, and translators), seven organisations, two schools (really one school with two campuses), one daycare, and one real-estate business, for a total of 18 publicly identifiable francophone establishments in Victoria.

[17] With church attendance severely on the decline in recent years – especially amongst young people (M. Robillard, *personal communication*, 24 December 2019) – the church likely plays little to no role in shaping G_2 speech. Furthermore, the parish in Victoria is practically LF-exclusive, so while it may well contribute to LF group-internal cohesion, it seems doubtful that it plays a role in HF/LF convergence.

Generally speaking, this means that G_1 speakers, on the whole, tend to interact with the broader francophone community far more rarely than G_2 speakers. While some may choose to involve themselves with the francophone community, there is no obligation for them to do so, and consequently dialect contact in French is very much optional for this group. Unlike G_2 speakers, who are likely to be introduced to and subsequently kept in close contact with the francophone community via their peer group from the school (see Section 2.4), contact with the francophone community is effectively optional for G_1 speakers. According to one participant who is well-versed in the history of the region, early G_1 arrivals *electively* isolated themselves as a means of protecting their francophone identity (14). This has a clear connection to a similar strategy adopted in early 20th century Quebec, termed *la survivance*, which consisted of "avoiding contamination by urban, English Montreal and maintain[ing] French-Catholic purity in the homogeneous environments of rural and small-town Quebec" (Levine 1990, 33).

(14) Les francophones qui sont arrivés dans les années quarante, soixante, ils– ils– je comprends qu'il fallait qu'ils se ferment sur eux pour s'identifier.

[VFC.LF_1.Béatriz.00/54/19]

'The francophones that arrived in the 40s, the 60s, they– they– I understand that they had to close in on themselves to identify themselves.'

In contrast, the vast majority of G_2 speakers attend francophone school from early childhood through to late adolescence[18]. Consequently, G_2 speakers' social networks are at least partially populated with their schoolmates – who are of francophone origins[19] even if they do not always choose French as their primary language of peer-level social interaction. Furthermore, even though the primary language of peer-level socialization for many G_2 speakers is English, the G_2 cohort is fully bilingual, and French-dominant G_2 speakers do choose to use French with their bilingual peers at least part of the time, including outside school grounds. This social network structure puts them into contact with children from families where a different dialect from their own is spoken, as well as with teachers or activity leaders who may hail from a wide variety of French-speaking regions, and therefore effectively *guarantees* dialect contact for G_2 speakers, as well as a social network where French is at least an option.

Many also participate in francophone youth group or other extracurricular activities, which has the effect of further flooding their immediate social network with other G_2 francophones. Outside the classroom, community organizations directed at francophone youth (such as the *Conseil jeunesse francophone de la Colombie-Britannique*, 'Conseil-jeunesse' or 'CJFCB'; see (15)–(16), (19)) play a similar role in French-language maintenance and dialect contact. Youth participation in these sociocultural activities is fairly high: the CJFCB offers around 35 different activities in the course of a calendar year, with average yearly participation of around 1200

[18] There is even a prize awarded at annual graduation ceremonies within the francophone school system for students who have attended a francophone institution from kindergarten to grade 12, possibly incentivising students to remain within the francophone system.

[19] In order to invoke their rights to French-language education under Article 23 of the Canadian Charter of Rights and Freedoms (the full text of which can be accessed on the Government of Canada webpage), parents must be francophone per the legal definition outlined in the charter.

individuals (which works out to around a quarter of BC francophone youth) [20]. Many participants claim that, through these youth-directed organizations or activities, they feel more connected to the francophone community (15), they note improvements to their confidence in French (16)–(17), or they observe that the social context of such activities is a better motivator to speak French (18) than the institutional context of school, where French is ascribed to authority figures like teachers (19).

(15) Ceux au Conseil-jeunesse, ils *hang out* avec des francophones plus que des anglophones. [VFC.LF$_2$.Jenna.02/03/34]

'The ones at *Conseil-jeunesse*, they hang out with francophones more than anglophones.'

(16) Avec le Conseil-jeunesse, t'y vas, tu veux pouvoir participer aux activités, tu veux, comme, améliorer ton français pour pouvoir être plus à l'aise et tout. [VFC.HF$_2$.Annique.01/21/17]

'With *Conseil-jeunesse*, you go, you want to participate in the activities, you want to, like, improve your French to be able to feel more at ease and everything.'

[20] Per CJFCB Annual Reports from 2017 on and Statistics Canada data from 2006.

(17) Je remarque que comme je di– j'ai m– un moins bon vocabulaire. Comme ça prend plus de temps à penser au mot que je veux mais là après, comme, disons une semaine pour les Jeux– mais ça c'était dix jours– il a, comme, une petite différence dans, comme, la confiance de mon français. [VFC.LF$_2$.Alex.01/43/26]

'I notice that like I sa– I have– a less good vocabulary. Like it takes more time to think of the word I want but then after, like, say one week for *Jeux*– but that was ten days– there's, like, a little difference in, like, the confidence of my French.'

(18) Quand il a une nécessité [sociale], je pense que c'est beaucoup plus facile que juste « ah tu dois parler français pour pouvoir faire ton essai de TDC [*Théorie de la connaissance*]. » [VFC.HF$_2$.Annique.01/21/25]

'When there's a [social] need, I think that it's much easier than just "ah you have to speak French to be able to do your *TDC* [Theory of Knowledge] essay."'

(19) Pour moi, comme, gran– avant– *pre*-Conseil-jeunesse, c'est quoi le français? C'est les enseignants. [VFC.LF$_2$.Jenna.02/04/03]

'For me, like, grow– before– pre-*Conseil-jeunesse*, what's French? It's teachers.'

2.3. Contact between HF- and LF-origin individuals

Victoria is home to mostly G$_1$ arrivals from other regions. Examination of provincial statistics[21] reveals that only 12% of BC francophones were actually born in BC, whereas 73% were born in

[21] Provincial statistics are deemed to be roughly representative of the situation in Victoria since it is home to the second-highest concentration of francophones in the province (after Vancouver).

another province, and an additional 15% were born in another country (Table 3.7.1; Chavez and Bouchard-Coulombe 2011).

Figure 2.1. Origins of provincial immigrants who claim French as their mother tongue within the BC francophone population [adapted from Chavez & Bouchard-Coloumbe (2011), Table 3.7.1]

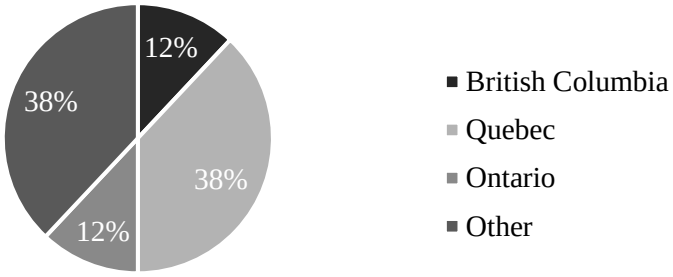


Table 2.1. Origins of French-speaking international immigrants within the BC francophone population [adapted from Chavez & Bouchard-Coulombe (2011), Table 3.7.2.2]

Country of origin	% International francophone population
France	26%
China	5%
Switzerland	4%
Belgium	4%
Romania	3%
Iran	3%
Hong-Kong	3%

Around half of all inter-provincial newcomers are from Laurentian French-speaking regions (Figure 2.1), with 38% from Quebec and another 12% from Ontario, whereas one-quarter of all international newcomers are from France (Table 2.1)[22].

This heterogeneity of origins likely contributes to the participant impression that G_2 speech is somehow 'mixed' (cf. (1)–(4)), but as noted in preceding sections, the community does not appear to provide many opportunities for this 'mixing' to occur. Perhaps exacerbated by the lack of speaker interaction overall (but also possibly because of prestige-based tension between speakers of HF and speakers of LF; see Section 1.2.3), there has been a historical lack of interaction specifically between speakers of different origins within the francophone community. From around the 1940s[23] to at least the 1990s, there was something of a rift between LF-origin speakers and HF-origin speakers (20)–(21), reflected in the presence of separate francophone organisations to serve each of these subgroups (*la Société francophone de Victoria* and *l'Alliance française*, respectively; (22)–(24)), but also in the (perceived or actual) socioeconomic class divide between the two communities (cf. (25)).

[22] Of course, it is not the case that *all* of these arrivals from elsewhere fall into the category of G_1 speakers by default, though it is highly likely that most are adults whose implicit usage norms would have been fully acquired in their region of origin, before immigration.

[23] The 1940s mark the decade in which there was a resurgence of public visibility for the francophone community, via efforts to form a club for new arrivals, *le Club canadien-français* (now, the *Société francophone de Victoria*), inaugurated in 1941 (Agnew et al. 1987).

(20) [Ils] m'ont pas souhaité la bienvenue, en disant, « ah bien t'es pas d'ici toi uh t'es pas– tu fais pas partie de la Socié– de la– la communauté francophone de souche, » et je me suis pas senti bienvenu puis j'ai dit « bon bien j'ai pas grand-chose à voir avec ces gens-là en fait. » [VFC.HF$_1$.Fernand.00/34/53]

'[They] didn't welcome me, saying "oh well you're not from here uh you're not– you don't belong to the Socie– to the– the native francophone community," and I didn't feel welcome so I said "well then I don't have much in common with these people, in fact."

(21) Ils avaient pas la même histoire que nous-autres, eux-autres, uh, ils venaient d'Europe. [VFC.LF$_1$.Donald.01/18/30]

'They didn't have the same history as us, they, uh, they came from Europe.'

(22) En fait je connais pas très bien les– les– les francophones d'ici vraiment, bien je connais– je connais les gens de l'Alliance-française certain. [VFC.HF$_1$.Hélène.00/05/13]

'Actually I don't really know the– the– the francophones from here really, well I know– I know the folks from *l'Alliance-française* of course.'

(23) Moi, tu-sais, au début quand je suis -t- arrivée ici : « Ah– oh– ah mais moi je vais pas à l'Alliance parce que c'est les français. » Tu sais je veux dire? … Ensuite– la Société-francophone avant c'était plus les québécois. [VFC.LF$_1$.Ida.00/36/38]

'Me, you-know, in the beginning when I arrived here, "Ah– oh– ah but I won't go to the *Alliance* because it's the French [from France]." You know what I mean? … Then– the *Société francophone* before was more the Quebecois.'

(24) Je pense qu'à un moment donné la– la Société-francophone s'est ouverte à tous les francophones plutôt de dire « non bien c'est– c'est quand même réservé aux gens qui sont d'origine canadienne-française. » [VFC.HF$_1$.Fernand.01/44/09]

'I think that at one point the– the Société-francophone opened itself to all francophones rather than say "no well it's– it's still reserved for people of French-Canadian origin."'

(25) C'est dispersé physiquement mais aussi il a des– aussi, par exemple, l'université, c'était plutôt des– des français de France qui étaient là, et puis qui avaient pas trop de rapport avec, par exemple, la base militaire. [VFC.HF$_1$.Fernand.00/50/12]

'It's physically dispersed but also there's– also, for example, the university, it was mainly French from France who were there, and who didn't have much to do with, for example, the military base[24].'

Furthermore, instead of new francophones arriving in Victoria and being welcomed by long-established francophones speaking some stable 'enclave' variety, they are far more likely to encounter another new francophone arrival (who may or may not share their patterns for speech), a perfectly bilingual francophone who uses English by default in the public sphere, or most likely of all, a monolingual anglophone. This situation also extends to the small number of G_1 speakers who have been established in Victoria for many years, and to the even smaller number of G_2 speakers who were raised as francophone in Victoria and have managed to maintain their French in adulthood.

[24] Since only Canadian citizens can join the military, it follows that the French-speaking population in the military is virtually *entirely* LF-origin.

The strong predominance of LF-origin individuals in Victoria makes it likely that – in the 'normal' circumstances of a stable community where speakers interact freely and across a range of contexts – G_2 individuals would converge on LF-oriented implicit usage norms. However, it should be noted that despite the important presence of LF speakers in the broader community, there is no guarantee that G_2 speakers (or G_1 speakers, for that matter) are actually in contact with the entire broader community (especially important to note, since much of the literature on convergence stresses the *face-to-face* nature of contact as crucial to the process), and therefore no guarantee that the linguistic exposure they receive reflects the proportional distribution of LF and HF speakers in the broader environment. In fact, within the bounds of the school, the staff tends to host a much higher concentration of HF individuals (around 15%[25]) than exists in the broader community (only around 4%). This suggests that the G_2 individuals who attend the school are exposed to more HF than their G_1 counterparts. In a similar vein, G_2 individuals who have at least one HF-origin parent receive proportionally much more HF exposure than a G_2 individual with no HF-origin parents. In other words, though LF predominates at the community level, there are different degrees of exposure to HF for different individuals within the community that might sway these speakers more towards HF (see Section 3.1.3).

2.4. G_2 French in the school

Generally speaking, G_2 speakers interact with their peers in English (though some do choose French, or vary between French and English), and confine their use of French (near-)exclusively

[25] Calculated on the basis of raw counts of staff that would have been employed at the school contemporaneously with my G_2 sample (from 1990–2010): 53 total individuals, eight of HF origins, 45 of LF origins.

to classroom settings and – to a far lesser extent – to the home. This effectively makes the school the emblematic bastion of French transmission in the public domain, as well as the primary locus of dialect contact for both G_1 and G_2 speakers (i.e., with teachers, with other students within the confines of the classroom, with parents of classmates, etc.). Brodeur acts to amalgamate G_2 speakers from all over the city, but also to encourage daily French language use (which may or may not be enforced in the home). In fact, for a large proportion of the G_2 population at large, school represents the only context in which French is consistently used within a community where it otherwise frequently loses ground to English (26).

(26) Quand on était à l'école, c'était, comme, *not a big deal* parce qu'on était entourés par le français de toute façon, mais maintenant je pense ma s– ma sœur et moi on a, comme– tu-sais on a dit à notre mère, *like* « fais un effort de nous parler en français parce que c'est vraiment notre seule chance. » [VFC.LF_2.Héloïse.00/57/04]

'When we were at school, it was, like, not a big deal because we were surrounded by French anyway, but now I think my s– my sister and I we, like– you-know we said to our mum, like "make an effort to speak to us in French because it's really our only chance."'

The importance of francophone schools within the G_2 population is also evident from provincial enrollment numbers. Chart 3.3.1 in Chavez and Bouchard-Coulombe's (2011, 20) report identifies approximately 4150 francophone individuals aged 5-19 (i.e., roughly 'school-aged') residing in British Columbia in 2006. The *Conseil scolaire francophone*'s annual report from the same year lists enrollment of just over 3800 students in francophone schools across the province. Thus, at least 92% (≈N = 3800/4150) of BC's school-aged francophones attend a francophone

school – and this is likely a conservative estimate since the figure from Chavez and Bouchard-Coulombe (2011) includes 18- and 19-year-olds who might no longer be enrolled in school. Furthermore, the CSF has also experienced the highest growth rate in the province compared to other school boards (Canada, Office of the Commissioner of Official Languages, and Consortia Development Group 2010). As such, most G_2 speakers across the province (Victoria included) must experience the full impact of their community's heterogeneity, as they are effectively forced into contact with other G_2 speakers, as well as teachers, administrative/support personnel, and parents of classmates, who may not all use the same variety of French.

This state of affairs differs markedly from that in other studies of language or dialect contact (e.g. Otheguy and Zentella [2012], or other studies of Spanish in the U.S.), where there is typically far less institutional support for minority languages than there is for French in Canada, and where the primary locus of language/dialect transmission is consequently the home or other informal domains rather than the institutional context of the school. The evaluation of G_2 speech as somehow connected to the school (e.g. '*Brodeurien*' from the example in (1)) or as 'more standard' than G_1 speech (less 'slang', fewer 'anglicisms', '*Radio-Canada*' accent) is echoed by many community members (27)–(30), which suggests that the school's important role in providing input to G_2 speakers may have had an effect on the linguistic structure of their speech.

(27) Je dirais que notre vocabulaire, notre français, est mieux que celle au Québec parce qu'on a moins de– de, comme, *slang*. [VFC.LF_2.Geneviève.01/34/29]

'I would say that our vocabulary, our French, is better than the one in Quebec because we have less– less, like, slang.'

(28) Mais je pense, comme, nous on a tellement fait un effort, surtout à l'école, de, comme, juste parler en français, que des fois je pense qu'on a moins d'anglicismes. [VFC.LF$_2$.Héloïse.01/18/24]

'But I think, like, we made such an effort, especially at school, to, like, just speak in French, that sometimes I think we have fewer anglicisms.'

(29) Ma mère elle dit toujours que– qu'on a l'accent Radio-Canada. [VFC.HF$_2$.Brigitte.01/10/35]

'My mum always says that– that we have the *Radio-Canada* accent.'

(30) Donc ils [G$_2$ speakers] ont– ils ont un français plutôt– un français plus standard, pas pi– pas typiquement québécois. [VFC.HF$_1$.Hélène.00/42/41]

'So they [G$_2$ speakers] have– they have a French that's rather– a more standard French, not pi– not typically Quebecois.'

Previous studies of French spoken by teachers in the classroom indicate that such speech contains a significantly higher incidence of more formal or standard speech forms, such as vastly inflated use of the negative particle *ne* (Poplack et al. [2015, 305]; see also Antes [2016]; Mougeon and Rehner [2019]). While of course teachers also employ implicit usage norms that do not necessarily conform to the prescriptive rules of school French (e.g., Poplack's [2015] findings for future temporal reference), the increase in features of school French (and concomitant decrease in those of vernacular French) constitutes a marked difference from the French that would be typically spoken in informal contexts such as the home. Since G$_2$ speakers in Victoria receive a significantly higher proportion of their French input from the classroom

compared to, say, speakers in a majority context who receive more input from a variety of formal and informal contexts, I will investigate whether G_2 usage norms are concomitantly affected by this exposure to school French.

2.5. Summary

Victoria differs from a conventional dialect contact situation in many ways. Its G_1 settlers are not permanent, but instead are rather ephemeral. The community is subject to a more or less continual cycle of loss and renewal as members assimilate or emigrate, and new arrivals come to take their place. The relative rarity of locally-born francophones, coupled with the near-constant influx of francophones from other parts of the country and the world, has created a rather heterogeneous and unstable demographic landscape in Victoria. This necessarily affects long-term residents of the region who hail from elsewhere (i.e., G_2 speakers), whose prolonged contact with dialects of French that are not their own may have resulted in a deviation from the speech norms they had upon arrival, but is particularly interesting to consider for speakers who would have *acquired* language in this context (i.e., G_2 speakers). Whereas French remains in the private domain for many G_1 individuals, which concomitantly lessens the likelihood that they will come into frequent contact with dialects other than their own, for G_2 individuals, community infrastructure represents an important source of diverse or heterogeneous language input, and also contributes to their language maintenance and the development of their francophone identity.

Many of the characteristics outlined in the foregoing sections paint a decidedly grim picture for the continued existence of French – let alone the normal development of shared sociolinguistic norms – within this community. Not only are francophones absolutely dwarfed by English, which in turn greatly reduces the regularity of francophone interaction, even long-

established francophones who *know* each other may not interact socially with any regularity, given the historical social divide between the two main dialect groups (see (20)–(25)). Furthermore, if most francophones in BC are arriving in the province and leaving again very shortly, this is certainly insufficient time to expand the existing G_2 population, but also might arguably be too short a time to serve as a linguistic model for established G_2 speakers. This leaves open the question of what G_2 speakers 'latch onto' linguistically from the diverse options offered up by established G_1 speakers. As an apparent (though as yet unconfirmed) result of this diverse input, G_2 speech is identified by the participants as a 'mix' of both main dialect groups in Victoria (HF and LF). Furthermore, the recognition of some distinct entity as 'G_2 speech' suggests that the second generation of Victoria francophones may have converged on shared norms for speech in spite of the many barriers to community cohesion as outlined above. The fact is, this community continues to produce G_2 (and, more rarely, G_3) speakers who are perfectly fluent speakers of French, which would indicate that the language is being acquired somehow, despite the many barriers that exist to impede its transmission.

Chapter 3. Data and method

I briefly mentioned in Section 1.3 that the method used for this study would have to allow me to a) determine whether G_2 speakers share implicit usage norms, and b) pinpoint, in their speech, which of these implicit patterns could be attributed to which of the three putative source varieties (HF, LF, or school French) – if any, since it is also possible that G_2 speakers have created norms 'from scratch'. These objectives can only be achieved using the COMPARATIVE method (Poplack and Meechan 1998; Poplack and Tagliamonte 2001) couched within the broader framework of VARIATION THEORY (Labov [1972; 1984]; Tagliamonte [2006; 2012]; Weinreich, Labov, and Herzog [1968] – to name but a few). At the very core of the variationist framework is the observation that variability is not random or accidental, but ubiquitous and inherent to language. Its methods function to elucidate whether the variability attested in G_2 speech can be associated, wholly or preferentially, with one of the provided G_1 models, or with the prescriptive rules of school French. Crucially, such variability is primarily observed in *spontaneous speech* – as opposed to, say, other types of language data (spoken or not) that have been rehearsed, planned, or otherwise premeditated, which typically contain far less variability. No such spontaneous speech data existed for Victoria French prior to this study[26]. Accordingly, I constructed a corpus

[26] Studies of French in British Columbia have been quite notably absent from the larger body of work on Canadian Frenches, though studies of French in Maillardville (a neighbourhood in Greater Vancouver, arguably the best-known historical enclave of French in BC) have picked up steam in recent years (e.g. Canac-Marquis and Guilbault 2014; Guilbault 2007; 2010; 2012). From personal communication with these scholars, the main difficulty in studying BC French seems to be in locating speakers – virtually impossible for outsiders without some connection to the local networks, but a barrier easily circumvented in the present study by my in-group status (see Section 3.1).

of spontaneous speech for the express purpose of this dissertation, which is described in Section 3.1. Prior to this study, the characteristics of French as it is spoken in Victoria were a completely unknown quantity. As such, my initial goal was to sample from as many different sectors of the community as was possible given the size and scarcity of the population, guided by my intimate knowledge of the community's social structure and history. The sample that I use in the present study has been refined somewhat to better reflect the axes of variability that are of interest here. Where my sample differs from the characteristics of the broader Victoria francophone community, I take care to explain why these slight deviations were necessary for analytical reasons.

The methodological procedure applied in this study differs slightly from that of a 'conventional' variationist analysis, since I focus on the relative importance of diagnostic factors to the selection of a variant, rather than provide a full model of the variability for each of the variables, which might include any number of *non*-diagnostic factors. During the initial stages of analysis, I also took care to examine the patterns of variation for each individual speaker – to confirm whether a pattern attested for the whole cohort was cohesively followed by the individuals within it. This was done to avoid subsuming different types of individuals under a single, researcher-imposed analytical category. After all, even though my divisions are well-motivated according to my knowledge of the social groups in the community, there is no guarantee that my participants will pattern according to the cohorts to which I have assigned them. In the Victoria situation, where the linguistic patterns are unknown, this was deemed an essential step to avoid erroneously conflating diverse patterns.

In Section 3.2, I outline some methodological specifics of this study, focusing on how I employ the comparative variationist method to exploit benchmarks in order to determine whether

convergence has taken place, and if so, in what direction and amongst which cohorts. I provide an overview of diagnostic conflict sites in Section 3.3. Finally, given the possibility that social factors may play a key role in shaping linguistic patterns in situations of dialect contact, I outline the factor groups included in this study to assess their importance in Section 3.4. These factor groups reflect the axes of variability within the social environment of my speakers that are likely to affect variant selection and are tested for each of the variables in turn.

3.1. Construction of the *Victoria French Corpus* (VFC)

Previous studies of convergence have tended to be limited by the type of data that is available to them. By this I mean that many are forced to rely on post-contact reconstruction of the conditions and end-products of contact from whatever contemporary evidence is available, and many studies can only examine the results of contact several generations after it has already occurred. I have the opposite problem for Victoria, since there is nothing from past generations to reconstruct. The ephemerality of the community with its loss-and-renewal cycle has precluded the establishment of any stable local variety, which means that I cannot compare the current state of G_1/G_2 speech in Victoria to any prior state of the language. On the flip side, it is possible to examine the situation in Victoria it as it unfolds synchronically, rather than having to reconstruct it from historical data, since I have access to speakers of different age groups and generations. This corpus serves an important purpose, since it gives us a rare birds-eye view of the effects of the multiple model environment *in progress.* Indeed, one future avenue of investigation for this community is to see whether the same patterns uncovered in this study have evolved with time or whether they are essentially 'on loop' as the community continues its cycle of demographic loss and renewal.

The absence of any spontaneous spoken data for Victoria French prior to this study made it necessary to first collect a corpus of sociolinguistic interviews that accesses participants' naturalistic vernacular. Corpora collected by non-community members can preclude participants from engaging in the most natural expression of their vernacular speech (cf. Poplack's [1978] finding that code-switching was used far less frequently, as well as quite differently, with out-group individuals than with in-group ones). As a core member of the Victoria francophone community who acquired French in Victoria from childhood, I was ideally placed to collect a corpus of naturalistic Victoria francophone speech. This corpus was collected over a period of 16 months from November 2016 to April 2018.

It should be noted that, although it is ideal for the interviewer to be an in-group member as far as data quantity and quality, it does raise the possibility of bias during the data collection procedure. This is a necessary risk in corpus linguistics, however, since the introduction of outsiders at the data collection stage causes speakers (especially those in minority communities, whose speech may be non-standard, stigmatised, etc.) to adapt their speech to better suit the interlocutor's speech – thereby factoring out any analytically interesting variability – or may even result in their failure to speak the minority language at all. Interviews by community members with community members yield a more reliably vernacular dataset, which is why the researcher being the main data collector is common practice in sociolinguistics. In some cases, researchers have trained an in-group community member to collect the data (for example, as Poplack and Tagliamonte [2001] did in their study of African Nova Scotian English), but ultimately, in my case, this simply replaces one bias (the researcher) with another (the interviewer).

3.1.1. The sociolinguistic interview

Although traditionally referred to as an 'interview,' the sociolinguistic interview is more of a semi-structured informal conversation, which makes it an ideal method to collect spontaneous speech data. The bulk of the participant-researcher interaction must be participant-led, with the researcher's primary goal being to elicit spontaneous speech, as exemplified by narratives of personal experience, without steering the conversation too forcefully. Vernacular features are thought to be quite stable, given that this is the variety first learned in childhood, and the narrative is well-known for being the conversation mode where vernacular features are most frequent (Labov 1984). A minimum of one hour of speech per participant is generally sufficient to obtain adequate amounts of data for analysis, but obviously longer interviews yield more data for variables that only surface very rarely, such as morphosyntactic variables. Interview guides of various question modules exist (e.g. the Q-GEN-II; Labov [1984]), but these are suggestions rather than absolute requirements. The interview session is generally concluded with a list of more structured questions that pertain specifically to participants' opinions and attitudes about language, about features that might characterize their own variety, or about their experience in a minority community. The idea is that this should elicit a more formal register, which is an important benchmark against which to assess the rest of the interview.

Although the aim at the outset was to strictly adhere to the basic sociolinguistic interview technique described above, it quickly became clear that this technique was designed to allow individuals who are not well-known to each other to become sufficiently acquainted or relaxed within a short amount of time, such that they may engage in vernacular exchanges within the time limits of the interaction. When an individual was relatively disconnected from my own social network (as was the case for only three of the participants), I did employ this technique.

However, because my own social network largely overlaps with that of the participants and I actually have years of personal history with many of them, the traditional line of questioning (from broad demographic questions to more specific questions about childhood, relationships, etc.) was almost never used, since the resulting conversation would have felt quite unnatural for both of us. Instead, this collection of interviews contains much reminiscence about shared childhood experiences, talk of mutual acquaintances or updates on activities since the last time we met, and informal chronicles about family history (including origins, when they arrived in Canada/Victoria, etc.). All interviews last an hour at a minimum, while some extend to over three hours.

Talk of language or personal experience as a minority-language speaker is peppered throughout these conversations rather than strictly confined to the end of the interview, as this is an extremely natural topic of conversation for minority-language individuals – particularly when talking *in* the minority language to another minority-language individual – but the end of the interview does contain a line of questioning about language specifically and nine of the later interviews employ a detailed language-background questionnaire[27]. The advantage of such an informal approach is that most interviews contain several stretches of spontaneous vernacular

[27] This questionnaire was slightly modified from the University of Ottawa ERP Laboratory's *Language Background Questionnaire – French-language long version* (Sabourin et al. 2016). It measures age of first exposure to a language, self-assessed language dominance and competence, length and depth of language immersion, and language used with a variety of interlocutors (including parents, siblings, friends, colleagues) in a variety of contexts (within and outside the home, school, work). This questionnaire was incorporated later in the data collection process, since content analysis of the language-directed portion of the interview provided sufficient information to broadly assess language usage patterns.

speech. A drawback is that full demographic or language-usage information for participants is not always directly recoverable from the interviews themselves, though they may be present in written form (e.g. participants' year of birth was noted on hard copies of their consent forms). Content analysis of the language-directed portion of the interview provided sufficient information to broadly assess language usage patterns for nearly all speakers, but occasionally, some specific information was unrecoverable[28].

There were several spontaneous group interviews where two participants were interviewed simultaneously in the same session. Often, I had arranged to meet one participant, and this participant brought along an additional person who also qualified for participation according to my initial criteria. My sample contains 11 such interviews: two of friends, two of siblings, three of married individuals, and four of a parent and their (adult) child. Group interviews maximize the likelihood of informal conversational modes, since the two participants generally engage in these speech modes with each other, even if they do not with the interviewer. The general exception to this is parent-child interviews, which may hinder the child participant from fully expressing themselves – perhaps by avoiding vernacular modes that may be uncomfortable for a parent's ears. However, in the present study, the concerns over parent-child interviews are

[28] It should be noted that this missing information does not pose any major problems for the interpretation of the results, since no major analytical categories (e.g. 'restriction', 'age', etc.) are missing for any of the speakers. For example, prior to the introduction of the language background questionnaire, information about minor details (such as whether a participant used French [exclusively or partially] with their siblings or what specific region of Quebec a particular participant was from) might not be available, but broader information about their use of French within certain contexts was readily available. Furthermore, the way that I constructed my index scores for these measures (see Section 3.4) ensured that the absence of these details was trivial, since scores were proportionally weighted to account for only those contexts for which information was available.

minimal. Of the four parent-child interviews, two were with underage or young-adult children (Iris, aged 14, and Angèle, aged 20) and two were with more mature adult children (Brigitte, aged 27 and Émilie, aged 53). Iris and Angèle's interviews were both excluded for independent reasons (see Section 3.1.2), and Brigitte and Émilie are well into adulthood themselves, with fewer concerns about censoring themselves in front of their parents.

3.1.2. Sampling

Given the purpose of this study (i.e., to assess whether G_2 speakers have converged on a shared usage model despite severe restriction to both frequency of French usage and domains of interaction), it was necessary to ensure that my sample contained individuals for whom sharedness was at least possible. The overall social structure of the Victoria community does not seem propitious to the development of shared norms, but impressionistic assessment indicates that this may have been achieved nonetheless. As such, I aimed to identify the sector of the population for whom development of shared norms would be most likely. If there is no evidence of sharedness within this sector, then it seems highly unlikely that such norms have actually developed. If this turns out to be the case, then the impressionistic judgments of cohesion provided by participants might instead reflect an idealistic, imagined community or social unity, rather than a genuinely *linguistic* one.

Since developing shared usage patterns relies on actually *using* the language in question to some degree, I deemed it necessary to focus on G_2 individuals who are past-and-present *users* of French – that is, those for whom French was the first language acquired in childhood (either sequentially or concurrently with English), who maintained French as the primary home language until adolescence, and who can presently sustain a comfortable conversation in French. This group of G_2 speakers can be contrasted with those for whom the school is the sole purveyor

of French input and the sole context of French usage. Many of the latter individuals still identify as francophones, but primarily for 'hereditary' reasons (i.e., their parents were raised in French), and they may not be able to sustain a conversation in French for any length of time. In contrast, the former individuals can be considered francophone in a *linguistic* sense as well as in a hereditary sense, being that they were raised in households where French was the dominant language during their formative years. This study – being a socio*linguistics* study – necessarily focuses on the group of 'linguistic' francophones, since it is not clear to what extent the individuals in the 'hereditary' francophone group can be considered 'full' (or even native) speakers of French[29]. G_1 individuals are uncontroversially native speakers of French, having spent the majority of their lives in French-speaking locales.

The 'linguistic' francophones would have had the opportunity to acquire French in Victoria under conditions as close to 'normal' as possible (i.e., via naturalistic transmission from parents to children in the course of regular interaction), despite the clear deviations from a 'normal' situation instantiated by the community context. Evidence from studies of L1 acquisition shows that children acquire complex patterns of internal and external conditioning well before the onset of adolescence (Roberts 1993), and evidence from dialect contact in children shows that convergent processes begin instantly and resolve quite quickly (within a few months of arrival, cf. Payne [1980]). Taken together, this suggests that G_2 children in Victoria who were raised in

[29] Some 'hereditary' francophones – i.e. those whose French was acquired virtually entirely within the school context – have French proficiency that impressionalistically seems on par with those of speakers who acquired French via 'traditional' transmission (i.e. directly from parents to children in the home). However, in order to avoid any potential issues caused by including speakers who might not fit the traditional definition of native speakers (and consequently may also have different linguistic patterns), the present analysis focuses on the 'linguistic' francophones only.

primarily French-speaking households would have been 'eligible' to participate in both the acquisition of complex systems and, possibly, convergence on systems unrepresented in their households. As a methodological safeguard against the problem of subsuming dissimilar speakers together under a single heading, initial stages of the analysis for each variable examined the individual trends present in the data, to verify whether speakers behaved cohesively within their cohort groupings. Furthermore, the analysis builds in measures of speakers who may be more restricted than others in order to directly test whether this has any effect on their participation in convergence.

In addition to participants who volunteered to participate after hearing about the study through friend-of-a-friend sampling (N = 16), targeted advertising in francophone association newsletters (N = 2), and in-person calls for participation at francophone gatherings (N = 7), I also recruited several individuals from my own immediate network of friends and acquaintances (N = 15) to ensure a sufficient number of individuals for whom French was the primary home language and who have the ability to sustain a conversation in French. Individuals who are not necessarily L1 speakers of French (e.g. those who were raised in predominantly or exclusively English-speaking households; N=6) were included in the initial data collection, generally because their spouse or parent had been invited to participate and they happened to be present during the interview, but these individuals were subsequently excluded from analysis. Another participant was excluded for technical reasons: the recording device malfunctioned, resulting in only 20 minutes of data, which is insufficient to evaluate most of the variables in this study. Two more participants were excluded because they had resided in Victoria for fewer than 10 years, which was (somewhat arbitrarily) deemed to be insufficient time for any implicit usage norms to stabilise. Given that newly arrived francophones may not immediately integrate into the

francophone community (for reasons of isolation, etc. as described in Chapter 2), I wanted to be sure to include only those participants who had had enough time to do so. A decade seemed like a reasonable time span for this to occur. This left a sample of 31 speakers for analysis.

3.1.3. Cohorts

The 31 speakers retained for analysis are distributed into cohorts according to their age of arrival (GENERATION) and the dialect that would have been present in the home in childhood (ORIGIN). Generation is determined on the basis of age of arrival in Victoria: G_1 speakers are defined here as those who arrived in Victoria after the age of twenty-five, whereas G_2 speakers are those who arrived in Victoria before the age of thirteen or were born there. Origin for G_1 speakers is coded according to the broad geographic area in which they were raised: HF speakers are those who were born and raised in France and LF speakers are those who were born and raised in Quebec[30]. All G_1 speakers were raised in stable, monolingual speech communities with either two HF parents or two LF parents. Origin for G_2 speakers is determined according to the dialect origin of their parents, who would have provided the early acquisition input for these individuals. In the G_2 group, 'HF-origin' speakers are those who received an important degree of HF input in early acquisition – that is, they have at least one parent who was born and raised in France. In the same vein, 'LF-origin' G_2 speakers are those who received LF input from their parents in early acquisition.

[30] Although Section 2.3 made note of the fact that speakers from multiple Laurentian-speaking regions reside in Victoria, my sample happens to contain only speakers from the province of Quebec – not unusual considering their relative predominance in the population at large (cf. Figure 2.1).

Table 3.1. Familial configurations of G_2 speakers

		HF-origin	LF-origin	Anglophone
HF-origin	Annique	Both parents		
	Brigitte	Father		Mother[a]
	Daniel	Mother	Father	
	Éric	Mother	Father	
	Hugo	Mother	Father	
	Frédéric	Father	Mother	
	Gilbert	Father	Mother	
LF-origin	André		Both parents	
	Isaac[b]		Both parents	
	Alex[b]		Both parents	
	Geneviève		Both parents	
	Joseph		Both parents	
	Dominique		Father	Mother[a]
	Jenna		Father	Mother[a]
	Fabienne		Mother	Father
	Héloïse		Mother	Father

NB: Dotted lines around two individuals indicate that they are siblings.

[a]: Though classed as anglophone on the basis of first language acquired in childhood, these parents are highly proficient speakers of French.

[b]: Isaac and Alex became step-siblings in late adolescence.

As shown in Table 3.1, this yields a number of different familial configurations ('HF' = HF+HF, HF+LF, HF+EN; 'LF' = LF+LF, LF+EN). Because most participants in the HF group actually have one LF parent (N = 5/7), it may be possible that they pattern more like LF speakers (cf.

Potowski's [(2011)] findings on *intra-familial* dialect contact[31]), but the categories 'HF' and 'LF' are intended to distinguish those who would have been exposed to HF in early acquisition from those who were not, under the assumption that this exposure may have contributed to differing linguistic patterns.

The determination of linguistic proficiency is notoriously difficult, but Table 3.1 can incidentally shed some light on the question of G_2 French proficiency, since it reveals that most of the G_2 participants in this study come from endogamous households; that is, households where both parents are francophone. Although all G_2 participants in this sample spoke French as the predominant language in the home, those from endogamous households have the added feature of having been unquestionably French-dominant for some portion of their childhoods – generally until the onset of formal instruction in English, around age 8. During their interviews, many participants recalled having been introduced to English in the school setting as young children (English lessons begin in grade three at Brodeur). It is also not the case that these individuals instantly 'converted' to English dominance once they acquired it as a second language, since they maintained French as the primary home language for the most part (see Section 3.4). These speakers are clearly different from 'heritage' speakers – that is, bilingual G_2 children who are no longer dominant in their mother tongue (i.e., the language of their G_1 parents), but instead are

[31] Although it could be argued that, within these familial configurations, the primary caregiver might be expected to have a greater influence on linguistic outcomes, this is unlikely to have any effect within this particular sample. With the exception of Frédéric and Gilbert (whose mother, Ida, a G_1 participant in this study, was a full-time parent), all G_2 speakers come from households where both parents worked outside the home full-time and therefore are likely to have shared parenting duties roughly equally. Consequently, primary caregiver does not seem like a relevant operationalisation for this sample, despite its apparent effect in similar situations of dialect contact elsewhere (Potowski 2011).

dominant in the language of the new community (cf. Benmamoun, Montrul, and Polinsky 2013) – since most continue to use French on a regular basis, if only in a restricted range of social contexts. The G_2 participants *not* from endogamous households can be thought of as equally proficient in French, since this was the primary (though not the only) language spoken at home in childhood through to adolescence. Furthermore, as we will see in Section 3.4, the five G_2 participants from exogamous households also maintained friend networks in childhood and adolescence where French was the preferred language of conversation, and continue to use French regularly with their parents in adulthood. In other words, although these speakers differ from monolinguals acquiring French in contexts where French is the majority language, they are also unlike 'heritage' speakers in terms of both manner and length of exposure.

The cross-tabulation of generation and origin across the entire sample yields four different cohorts (Table 3.2), labelled here and throughout with the combination of their origin and generation number (i.e., the cohort of first-generation individuals who have Laurentian origins is labelled the LF_1 cohort, and so on). One participant defied straightforward classification at this stage: she was born in Saskatoon to two HF-origin G_1 parents, raised there speaking French, and then moved to Victoria as an adult. She is unlike the HF_1 cohort in that she did not acquire an HF variety in a stable, monolingual HF locale and then import that variety to Victoria, and she is unlike the HF_2 cohort in that she was not raised in Victoria and therefore unlikely to pattern with the rest of the G_2 speakers. Rather than arbitrarily shoehorn this participant into one cohort or the other, which would skew the results, I deemed it more prudent to exclude her altogether, leaving a total of 30 speakers for analysis (Table 3.2).

Table 3.2. Number of participants, by cohort

N participants	HF	LF	Total
G_1	HF_1 = 4	LF_1 = 10	14
G_2	HF_2 = 7	LF_2 = 9	16
Total	11	19	30

The internal linguistic patterns for each of the variables, briefly described in Section 3.3, are analysed individually for each cohort, under the assumption that each may pattern differently. However, these divisions can also be broken down into their component parts (origin and generation) to be used for the comparative analysis of patterns. This will help in disambiguating whether speakers' different origins are responsible for differences in linguistic patterning (as the case may be), or whether it is generation that has a greater effect on patterns. It may be the case, for example, that differences in patterns occur along the lines of origins – that is, a pattern that is shared by the HF speakers only (of both generations) would reveal their lack of convergence with LF-origin speakers. In a similar vein, a pattern that is attributable to generation – for example, one that is shared by G_2 speakers where G_1 speakers have distinct patterns – would indicate that convergence-type processes are underway in G_2. Such findings are crucial to the research objectives, which aim to uncover whether G_2 speakers share patterns, and if so, which linguistic model(s) in their environment would have contributed to these shared patterns.

3.2. Comparative benchmarks

Fundamentally, the comparative variationist method allows the analyst to objectively compare patterns – be it between two or more cohorts of speakers of the same variety (Poplack 2015), two or more different languages (e.g. French, Spanish, Portuguese, and Italian; Poplack et al. [2018]),

the same language at different points in time (Poplack and St-Amand 2007), languages in contact (Poplack and Levey 2010; Poplack and Meechan 1998), or different dialects/varieties of a single language (Digesto 2019; Kastronic 2016), the latter being the specific axis of comparison used in this study. This method of analysis will enable me to isolate any shared patterns in the speech of G_2 speakers, as well as determine whether these patterns align with those from any of the purported source varieties present in their linguistic environment (HF, LF, or school French).

In order to determine whether changes to a given variety have been induced by contact, the ideal method would be to first measure the relevant phenomena in the speech of individuals *pre*-contact, that is prior to their relocation to the context in which they will come into contact with a dialect different from their own, then again post-contact (Britain 2012). This is not a set of circumstances that is easy to come by. However, the next best thing, as proposed by Poplack and Levey (2010) is to locate (preferably contemporaneous) benchmarks of the same varieties that have *not* undergone the same kind of contact, and to compare one's data to that. To exemplify with one of the variables included in the present study, if non-contact LF has categorical assibilation, and non-contact HF does not, the presence of categorical assibilation in G_2 speech, where LF and HF are in contact, would suggest that G_2 speakers are converging on LF, not HF.

In this study, I test three possible sources of influence on G_2 French (LF usage, HF usage, and school French, defined below) and thus need at least three benchmarks to effect the comparisons described above. The benchmark data for usage norms in HF and LF comes from previous studies on each of these variables – with priority given to *variationist* studies if they exist – but since the rates and underlying conditioning differs according to each of the variables I have elected to examine here, the full presentation of LF and HF benchmark data is reserved for each individual analysis chapter. In addition, I also include as part of the usage benchmarks any findings from

variationist studies that have established a link between a variant's occurrence and factors such as socioeconomic class, level of education, or speech style, under the assumption that variants that are sensitive to these factors reflect their evaluation as prestigious or stigmatised. For instance, a variant that is strongly disfavoured in formal speech styles is more likely to be considered stigmatised, whereas a variant that occurs at a significantly higher rate in individuals of high social standing or with advanced education is more likely to be considered prestigious.

Two general issues are introduced by these comparative benchmarks. Firstly, although presented here as fundamentally internally consistent, there is a certain amount of regional variation in both LF and HF. Even features that are considered emblematic of a particular variety (e.g. assibilation in LF) are not *universally* employed across every instantiation of this variety. However, I was careful to select variables that were largely consistent within each of the varieties in question and, importantly, I privileged variables that are also *emblematic* in some way. Assibilation might not be universal across *every* variety of Canadian French, but many studies have found it to be near-categorical, and it is often referenced as an emblematic feature of this variety. Similar statements could be made about the other variables in this study as well (with the possible exception of yes/no questions, which is generally understudied compared to the other three). This high level of internal consistency was deemed acceptable for the purposes of the present study.

The second possible limitation of the benchmarks is that the type of data obtained in benchmark studies is not always comparable, one-to-one, with the spontaneous corpus data of the present study – especially for HF, since variationist research is far less available in France as compared to Canada. For example, some of the studies I used to obtain an idea of the rates expected for the yes/no question variable are very old (dating back to the 1960s in some cases),

or they consist of examinations of plays (Terry 1970), radio speech (Behnstedt 1973), or personal correspondence (Pohl 1965) – none of which corresponds perfectly to the spontaneous, contemporary speech collected for the present study. This issue is not unfamiliar to sociolinguists, who are often required to 'make the best use of bad data' (Labov 1994, 11). While the scope of the present study precludes this issue from being rectified (since it would require collecting vast corpora of benchmark data, in addition to the Victoria corpus), it is a limitation that must be acknowledged nonetheless.

Lastly, the school French benchmark is compiled from several sources. Firstly, I made note of prescriptive rules for each of my variables, within BC Ministry of Education curriculum documents and pedagogical guides, and within prescriptive grammars (mainly the 16th edition of *Le bon usage*; [Grevisse and Goosse 2016]). I also referred to the findings of the *Recueil historique des grammaires du français* (Poplack et al. 2015), which summarizes the fluctuations and inconsistencies in grammatical prescription over the last few centuries for a large number of variables, including *avoir/être* and future temporal reference. Finally, I incorporate findings from Poplack's (2015) study of teacher's speech in order to establish whether the rules contained in prescriptive sources are actually used in teachers' praxis (where applicable, given the variables under study). I also reviewed studies from the scientific literature on pedagogy which assess the degree to which pedagogical objectives and materials correspond to 'real' speech (e.g. Antes 2016; Etienne and Sax 2009). The prescriptive rules are different for each of the variables, and so the full description of the school French benchmark is reserved for their respective chapters.

Note that, in each of these sources, the prescribed rules are assumed to be uniform, in the sense that contemporary grammatical manuals are not written differently for HF and for LF[32].

Comparisons in comparative variationist work are generally achieved using three main lines of evidence in each of the quantitative patterns under investigation, including a) which factor groups (i.e., operationalizations of hypotheses regarding what affects variant selection) are selected as significant by the multivariate model, b) the relative importance of these factor groups vis-à-vis each other (= the MAGNITUDE OF EFFECT, ordered according to the *range*), and c) the ordering of factors within a factor group (= the direction of effect, or CONSTRAINT HIERARCHY). As the discussion in Poplack and Tagliamonte (2001, 91) states, certain factor groups might not achieve significance in a given analysis *not* because they do not affect variant selection, but because the dataset is too small for the model to select them as significant. As such, their recommendation is to refer mainly to the constraint hierarchy to draw conclusions about whether a given phenomenon patterns in parallel across a number of different benchmarks. For instance, if two cohorts share the same set of significant factor groups, but the constraint hierarchy within these factor groups differs from cohort to cohort, then the implicit usage patterns for these cohorts cannot be said to be aligned (see also Grevisse and Goosse [2016, 731–32]).

This method has been used to great effect in the broad field of contact linguistics. Because it relies on objective comparison of underlying patterns, it allows the analyst to determine more confidently whether a suspected change is actually due to contact, or perhaps instead due to internal evolution, linguistic universals, coincidence, etc. Poplack and Meechan (1998), as well

[32] Though it seems clear that many prescriptive sources (e.g. Grevisse and Goosse 2016) are written from the perspective of HF, as I demonstrate in the relevant chapters.

as Poplack and Tagliamonte (2001), Poplack and Levey (2010), and Poplack (2018), convincingly demonstrate how reliable this method is in determining, among putative sources for a given phenomenon, which source(s) contributed to the end product. They achieve this using multiple axes of comparison to rule out alternative explanations for the development of a contemporary form, a method also used in this study, as well as a number of diagnostic variables. The next section outlines in more detail how each of the variables selected for this study meet the criterion of distinct patterning in each of the putative source varieties.

3.3. Diagnostic variables

Many studies of dialect contact (in established communities and in new territory) focus on phonetic, phonological, and/or lexical variability, which all have the advantage of being highly salient and labile, with little to no overlap in production[33]. In the case of lexical variation, this consists of wholly substituting one expression for another; for example, 'dustbin' in Southern England English for Canadian English 'garbage can' (cf. Chambers 1992, 677). Determining whether a sample of Canadian English speakers has converged on Southern England English (at least as far as the superficial lexical level) would therefore be as easy as counting up the number of times a speaker uses 'dustbin' and the number of times they use 'garbage can'. Other areas of the grammar are far more challenging to assess.

In phonetic or phonological variation, the allophones or phonemes under study must be situated at different enough points in acoustic space such that they may be considered categorical – keeping in mind that there is certainly some within-category vaccilation in spontaneous

[33] Especially if the data is obtained using wordlists or elicitation rather than spontaneous speech, which is a frequent data collection method in studies of dialect contact.

production[34]. The term 'mixture', often used in dialect contact research implies that two distinct varieties have each contributed elements of their systems to the 'mixed' system, or perhaps that speakers with 'mixed' systems alternate between options provided by each of the source varieties. In the abstract, this might appear straightforward to determine, but in reality, it is quite difficult to operationalize empirically: how does one decide whether the natural vagaries of acoustic production are 'mixed' systems or simple instantiations of inherent variability (which characterises every level of speech)? Analysts in the broad field of dialect contact often speak of allophone *x* belonging to one dialect, and allophone *y* belonging to another, yet they infrequently account for the natural variability of individual productions that must be subsumed by each of these allophones. The risk with this method is that there might be overlap (great or little) between the individual productions of these allophones in speech that could confidently be classed as *un*mixed. That is, speakers who have never been in contact with 'allophone *y*' dialects might still produce iterations of allophone *x* that fall within the range of allophone *y*. Therefore, when examining individual speakers for evidence of 'mixture', it might be the case that the analyst misinterprets a variable realization of allophone *x* within the acoustic bounds of allophone *y* as a 'mixed' system, when in fact it is an instantiation of inherent, production-based variability in allophone *x*.

To give a concrete example, take the vowel data shown in Kerswill and Williams (2000, 88–89). They show that mothers raised outside but currently residing in the 'New Town' of Milton

[34] Poplack et al.'s (2020) study of phonetic integration of borrowed material provides a clear demonstration of how to accomplish this.

Keynes realize the MOUTH vowel (au) variably, as the allophones [ɛː], [aːə], [æʊ], or [aʊ][35]. Their findings suggest that East London mothers tend to prefer the first two variants, whereas Home Counties mothers prefer the latter two, and West and North London mothers are 'intermediate' between the two preferences. Here, we run into the first issue. The 'intermediate' mothers are not all *individually* intermediate: some seem to prefer [ɛː] or [aːə], whereas others prefer [æʊ] or [aʊ]; it is not the case that each West and North London mother uses 'intermediate' rates of each allophone. In other words, some West and North London mothers pattern like Home Counties mothers (in that they prefer [æʊ] or [aʊ]), and others pattern like East London mothers (in that they prefer [ɛː] or [aːə]). Kerswill and Williams use the data from these three sets of mothers as the benchmark against which to compare the vowel productions of mothers and adolescents raised in Milton Keynes, who both strongly prefer the variant [aʊ] (which, incidentally, happens to be the variant associated with Received Pronunciation).

The adolescents use other variants in addition to [aʊ], which is interpreted as an effect of exposure to the vowel systems of the mothers from outside Milton Keynes, and therefore a 'mixed' system. However, recall that there was considerable individual variability to be found in the productions of West and North London mothers. To what extent can Milton Keynes adolescents be confidently identified as 'mixed', when West and North London mothers display similar 'mixture' without having been subject to the same contact conditions as Milton Keynes adolescents? That is, two completely different linguistic environments (no contact: North and West London versus contact: Milton Keynes) and populations (mothers versus adolescents) seem to yield the same linguistic trends. It therefore seems questionable to conclude that the latter is

[35] Note also that these are only the allophonic realizations that the authors "felt [they] could reasonably classify" (Kerswill and Williams 2000, 87) – there are certainly more.

the result of 'mixture' without considering the same possibility for the former. In other words, the diagnosticity of phonetic or phonological variables depends on the one-to-one association of a given phoneme with a given dialect – as well as convincing demonstrations that this is the case (e.g. by incorporating analysis of benchmark productions by native speakers of the dialect in question alongside those of the new arrivals[36], as in Chambers [1992]).

In contrast, morphosyntactic or syntactic variants cannot generally be associated one-to-one with a given dialect – each of the dialects being compared uses the same variants. Thus, determining what counts as 'mixture' becomes even more difficult, because two very different dialects could use exactly the same morphosyntactic forms at different rates, with contrasting patterns of underlying conditioning (described and exemplified in Section 1.3), or under different sociostylistic circumstances. Since in many cases, the forms involved in a dialect contact situation belong to both dialects in question, it is unclear how a notion like 'mixture' could even be assessed for morphosyntactic or syntactic variables. Would speakers have to show evidence of each and every underlying constraint attested in benchmark data for their variety to be considered 'unmixed'? Conversely, if speakers were missing a constraint or had incorporated one or more different constraints, could this unambiguously be considered evidence of 'mixture'? Such fluctuations in constraint hierarchy could just as easily be due to internal evolution or changes in progress, and may not have anything to do with contact at all. These

[36] As Siegel (2010, 23) points out, "inconsistent use of a variant may reflect variation rather than incomplete acquisition – especially when dealing with closely related speech varieties." In other words, using appropriate comparative benchmarks is an *essential* step in order to avoid mislabeling garden-variety inherent variability as genuine change to the speech of new arrivals (see also Section 3.2).

problematic questions are resolved in the present study by establishing adequate benchmarks for comparison.

For the purposes of determining whether G_2 usage aligns with LF usage, HF usage, or school French, I refer to Poplack and Tagliamonte's (2001, 100) PRINCIPLE OF DIAGNOSTICITY, which stipulates that variables used to make such determinations must be *diagnostic*: uniquely associated with one source variety but not another. As we briefly saw in Table 1.1, some of the variables I selected permit comparison between all three source varieties (HF usage vs. LF usage vs. school French), whereas others only test HF usage vs. LF usage, or usage vs. school French. However, each of the variables selected here is considered to be diagnostic because it has been shown in previous work to have patterns that clearly differ in the usage norms of HF, of LF, and in the rules of school French. These characteristics are reiterated in each of the analysis chapters, as well as which specific elements of the underlying conditioning are considered to be diagnostic, but are summarized briefly here.

The linguistic variables I have selected represent four different and well-established linguistic patterns in their respective varieties, which also feature differing degrees of frequency (occurrence of the variable context within the entire corpus), salience (variants that are emblematic of one variety but not another, or otherwise socially significant), and explicit commentary by the normative enterprise. I hypothesize that these characteristics might make the underlying patterns for some variables more obvious, or may perhaps act to reinforce implicit usage norms for some variants. For example, a variable that is not particularly frequent and that does not feature any particularly socially significant (and therefore salient) variants will – hypothetically – be a more challenging pattern for G_2 speakers to make out than one that is highly frequent with either a very salient, emblematic variant or one that is either mandated by

grammatical prescription (and therefore likely explicitly taught in school) or otherwise normatively valued.

3.3.1. Assibilation of coronal plosives

Assibilation is the phonological process by which coronal stops become sibilant affricates before high-front vowels or glides: that is, /t/ becomes [ts] before /y, i, j, ɥ/ (‘tu’: /ty/ → [tˢy]). There are three possible variants, illustrated in (31)–(33) with examples from a single speaker in identical phonological contexts.

(31) J’ai trouvé que c’était vraiment le même genre d’**architec<u>tu</u>re** [aʁʃitɛk**<u>ty</u>**ʁ].
[VFC.HF$_1$.Julianne.00/32/01]
‘I found that it was really the same kind of architecture.’

(32) Il avait fait une **pein<u>tu</u>re** [pɑ̃**<u>tʃy</u>**ʁ] impressionniste. [VFC.HF$_1$.Julianne.00/40/22]
‘He had done an impressionist painting.’

(33) En fait ils ont acheté une **voi<u>tu</u>re** [vwa**<u>tˢy</u>**ʁ]. [VFC.HF$_1$.Julianne.00/53/59]
‘In fact they bought a car.’

Crucially for my purposes, assibilation is both highly salient and emblematic of LF, whereas plosives are by far the dominant articulation in HF varieties. Furthermore, though some speakers of HF varieties have been found to assibilate and/or affricate infrequently in other studies (e.g. Junová 2018), these articulations are stigmatised and confined to a very narrow sector of the population. In contrast, even formal broadcast LF speech – that of *Radio-Canada* announcers –

features categorical assibilation (Bigot 2008). Given the stability and uniformity of assibilation (or lack thereof) in each of the benchmarks, this variable is ideally suited to our purpose of determining whether G_2 francophones align themselves with HF or with LF usage patterns. Note, however, that prescriptive sources remain more or less silent on this pronunciation – if they mention it at all, it is simply to say that assibilation is associated with Canadian speakers, but there is no prescriptive rule that prohibits or discourages its use. This variable is also interpretable solely from the rates, given its categoricity in the LF benchmark. Since the primary angle of investigation in this study depends on the analysis of underlying patterns, and such patterns are not discernible from categorical, phonologized phenomena, the remaining variables are all (morpho)syntactic.

3.3.2. Avoir/être

This study targets a somewhat simplified paradigm of auxiliary selection (compared to, say, that of Sankoff and Thibault [1977] or Rea [2016]); the focus here is on the variable use of the auxiliary *avoir* in contexts where prescriptive injunction would mandate the use of *être* (i.e., the 'Vandertramp' verbs) (34)–(35):

(34) Il en avait neuf qui **ont**$_{[AVOIR]}$ vraiment **resté** de la maternelle jusqu'à la douzième. [VFC.LF$_2$.Isaac.00/06/29]

'There were nine who actually stayed from kindergarten until twelfth [grade].'

(35) Quand je **suis**$_{[ÊTRE]}$ **resté** uh à Vancouver pendant deux ans, c'était genre, Queensborough. [VFC.LF$_2$.Isaac.00/31/32]

'When I stayed ['lived'] uh in Vancouver for two years, it was like, Queensborough.'

There is an identifiable implicit usage norm for LF (*avoir* is acceptable with certain verbs but unacceptable with others), but in this case the implicit usage norm clashes with the prescriptive rule for this variable (never use *avoir* with these verbs). In other words, this is one variable where speakers might rely on what they learned in school to guide their variant selection, since one variant (*avoir*) is categorically *pro*scribed by the normative enterprise in these contexts, which makes it a salient form to avoid. This variable is therefore quite a useful test case of whether G_2 speakers refer to prescriptive injunction or normative evaluation to shape their patterns – as opposed to implicit usage norms. It is less useful as a diagnostic *between* LF and HF, since what little information exists on *avoir/être* in HF suggests that it patterns in much the same way as in LF. However, the use of multiple variables in this study prevents this from being too problematic: assibilation is invoked to distinguish between HF and LF, *avoir/être* is invoked to distinguish between (HF/LF) usage and prescriptive rules, and the other two variables (described in the next sections) test the effects of all three putative sources at once.

3.3.3. Future temporal reference

There are three competing variant forms to express future time in French, two of which occur in sufficient numbers in spontaneous speech to be analysed in this study[37] (36)–(37):

[37] The third variant (the *futurate present*, e.g. "je <u>prends</u> trois cours au mois d'avril") is so rare that it is often excluded in other variationist analyses (Grimm and Nadasdi 2011, 174; Poplack and Dion 2009, 572). It was similarly scarce in the present study and was consequently excluded.

(36) Periphrastic future: Il lui reste trois cours, il **va** en **prendre**[PF] un au mois d'avril. [VFC.LF$_1$.Albert.00/51/47]

'He has three courses left, he's going to take one in the month of April.'

(37) Synthetic future: Il dit, « je **prendrai**[SF] mes cliques puis mes claques puis j'**irai**[SF] faire d'autre chose de ma vie. » [VFC.LF$_1$.Albert.01/49/43]

'He says, "I'll pack up my stuff [lit. take my clicks and my clacks] and I'll go make something else of my life."'

Previous work on future temporal reference reveals two main differences between HF and LF: not only does the synthetic variant figure more prominently in HF, it has a broader context of occurrence than in LF, where the synthetic variant is both proportionally rarer and largely confined to negative polarity contexts ((38); Poplack and Dion 2009).

(38) Quand Benoît **sera**[SF] plus là– qui ils **vont trouver**[PF], je sais pas. [VFC.LF$_1$.Albert.00/10/06]

'When Benoît isn't there anymore– who they'll find, I don't know.'

The polarity constraint is also operative in HF, but to a *much* weaker degree (Roberts 2012), whereas adverbial specification shows some evidence of being ranked higher than polarity (Gudmestad et al. 2020). Other constraints have been tested (including grammatical person, lexical verb, and temporal distance), but none were selected as significant. This contrasts with LF, where conditioning is much richer overall, and where several studies have consistently found

polarity to be a virtually categorical constraint on future temporal reference. This variant also has a fairly consistent prescriptive rule (the periphrastic future should be used for events that are to occur in the near future), which not adhered to in both LF and HF usage. In addition, the synthetic future is favoured by formal speech style in LF (Poplack and Dion 2009) and by speakers with higher educational attainment in HF (Roberts 2012). Note the contrast between the prescriptive rule and the sociostylistic evaluation: nowhere does the prescriptive rule stipulate that the synthetic future is preferred, but it has nonetheless garnered a positive sociostylistic evaluation in the LF usage norm.

Thus, future temporal reference represents an ideal three-way test of putative influences on G_2 speech: G_2 speakers cannot learn the polarity constraint in the classroom, nor can they learn the temporal distance rule from usage norms alone. Multivariate analysis allows the effects of adverbial specification (HF) to compete against those of polarity (LF) and temporal distance (school-F), meaning that it should be obvious which of these influences has had any effect on G_2 speech.

3.3.4. Yes/no questions

There are four ways of asking yes/no questions in French: inversion (39), the *-tu* particle (40), the interrogative phrase *est-ce que* (41), and rising intonation (42):

(39) « **Pouvez-vous**$_{[INV]}$ élaborer? » Moi je m'adresse à lui qu'en anglais.

[VFC.LF$_1$.Catherine.01/57/20]

'"Can you elaborate?" I'm addressing him only in English.'

(40) **Tu restais-tu**[tu] chez– chez uh Marie? [VFC.LF1.Catherine.01/49/42]

'Did you live at– at uh Marie's?'

(41) Tu-sais, **est-ce que**[ECQ] je suis obligé d'aimer ma mère? [VFC.LF1.Catherine.02/08/21]

'You-know, am I obligated to love my mother?'

(42) Le collège-Pearson ici, **tu connais**[INT]? [VFC.LF1.Catherine.01/43/40]

'Pearson-College here, you know it?'

The differences between HF and LF usage in yes/no questions are robust and long-standing, having been well-documented in diachronic data (Elsig 2009) as well as in many synchronic studies. The HF system is largely dominated by the intonation variant, whereas the LF system uses all four variants to some degree. Within the latter system, each variant has a specific function or preferred context of occurrence (e.g. intonation is preferred in negative polarity contexts; *est-ce que* is used as a hyperstyle marker; inversion categorically occurs with 2P subjects, etc.; cf. (Elsig and Poplack 2006)). The prescriptive rule for this variable also differs from both LF and HF, in that it strongly recommends variants that are rare in usage in both varieties (*est-ce que* for LF, inversion for HF), and stipulates that the intonation variant be categorically confined to informal contexts – which is not the case in HF nor in LF. Finally, there are sociostylistic associations with some of the variants in this system, which differ between HF and LF: the -TU particle, for instance, is strongly associated with informal LF speech but possibly stigmatised in HF speech, whereas intonation is attributed to formal style in LF, but informal style in HF.

These differences make this variable useful for the purposes of this study, but the complexity of the underlying systems (particularly for LF, which makes use of all four variants) compared to those for future temporal reference and *avoir/être* (which both have binary variants and fewer diagnostics to handle) makes this a particularly rich area of investigation. It might be the case that G_2 speakers have no problem aligning themselves with one of the putative sources for assibilation, *avoir/être*, or even future temporal reference with its near-categorical polarity constraint, but that alignment becomes considerably more challenging when faced with a variable where each of the usage norms, as well as the prescriptive rule and the sociostylistic evaluations assigned to some variants but not others, is so complex.

3.4. Extralinguistic factors

As the reader will recall from Section 1.2, there is some disagreement within the dialect contact literature as to whether outcomes of convergence are entirely down to frequency of a given form or predominance of a particular dialect in a mixed community, or whether these outcomes are directly affected by social factors such as status of the speakers involved in the dialect contact. Trudgill, for example, is a staunch defender of the strongly deterministic stance:

> "I maintain that, given sufficient linguistic information about the dialects which contribute to a mixture, and given sufficient demographic information about the proportions of speakers of the different dialects, it is possible, within certain limitations, to make predictions about what the outcome of the mixture will be, at least in broad outline." (Trudgill 2004, 26)

Others have countered this stance with evidence that suggests an 'overruling' effect of social motivation, perhaps to avoid certain stigmatised features – as was the case in Otheguy and Zentella's (2012) study of Caribbean and Mainland Spanish speakers in New York City, where converging speakers avoided the variant used by the majority of the population because this variant is stigmatised within the broader hispanophone community. Although this issue does not form the primary goal of this dissertation, the community studied here provides us with an opportunity to examine it at least in passing, since it involves contact between cohorts that are not equally represented within the community (LF speakers vastly outnumber HF speakers) and the choice between these two dialects has social significance within the Canadian context (cf. Section 1.2.3).

Some studies have also uncovered significant effects for factors pertaining to degree of connection to the community (e.g. those summarised in Dodsworth [2017]). The Victoria situation directly invites questions about the degree of (face-to-face) contact required for convergence-type outcomes to occur, and so I operationalised a number of such factors as well. It should be noted that the manner in which an analyst decides to operationalise a factor group reflects their hypothesis about what affects variant selection, and that different operationalisations of effectively the same hypothesis can yield different results[38]. Rather than impose a single analytical perspective on the data of this as-yet-unknown community, I decided at the outset to examine the same question from multiple angles by operationalising the same hypothesis in several different ways.

[38] Walker's (2012) study of different ways to operationalise 'frequency' is a prime example of this concept.

For instance, connection to the community is hypothesised to have some important effect on the outcome of dialect contact. But should this be measured using some assessment of social network, for example as Payne (1980) did, by operationalising the number of times children were mentioned by their peers? Or should it instead be calculated on the basis of an individual's interaction with the infrastructures provided by the community (and, presumably, the people contained within them)? To an extent, this hypothesis is also quantifiable by coding for the range of contexts in which speakers claim to use French (e.g. outside the home vs. within the home). I decided to incorporate as many different operationalisations as were possible from the data available to me, in order to more fully explore the possibility of extralinguistic effects on convergence outcomes. I will emphasize at this point that these factor groups are not orthogonal with each other – which does not preclude them from analysis, but non-orthogonal factors cannot be included together in the same multivariate model. If two factor groups that purport to measure different things actually measure the *same* thing, including both in a multivariate model would yield errant results. As such, the most straightforward way to extract information from these factor groups without drawing erroneous conclusions is to evaluate them with bivariate analysis, which allows the analyst to determine whether a given factor group has a statistically significant effect on variant selection. I perform this using tests that return an exact *p* value wherever possible (e.g. Fisher's two-tailed exact tests, Goldvarb bivariate runs), with the significance threshold set at $p < .05$.

3.4.1. Contextual restriction

The first of my operationalisations consists of a measure that taps into the contexts in which speakers use French. In the Victoria situation, such a measure is especially important because previous research has suggested that speakers who use French in a restricted range of social

contexts often lack the vernacular features that characterize their less restricted counterparts (e.g. Mougeon and Nadasdi 1998; Sammons, Nadasdi, and Mougeon 2015). Certain sectors of Victoria's broader G_2 population are similar to the participants in Mougeon and Beniak's (1989) study, in the sense that they have "received most of their exposure to French in the formal domain of the school" and therefore might "lack familiarity with various non-standard features" (*ibid*. p. 5). Generally speaking, I have observed that the G_2 speakers who do not speak French as a matter of daily routine in their homes are those who are most likely to prefer English for socialisation. G_2 individuals who use French at home (especially those who use French with their parents *and* their similarly-aged siblings) are far more likely to opt for French in their peer-group interactions than G_2 individuals for whom English is the primary home language.

While the G_2 speakers of the sample isolated for the present study distinguish themselves from these other sectors of the G_2 population by their more frequent use of French and their exposure to French in domains other than the school, there are nonetheless certain intra-sample differences along the lines of contextual restriction. To quantify these differences, I adapted elements of Mougeon and Beniak's (1991) restriction index, which attributes an overall score to each individual based on their responses to 11 different language-use questions in order to determine the range of contexts in which each participant has the opportunity to use French. I extracted information on language-use restriction through content analysis and systematic questionnaire data where the latter was available (see Section 3.1.1). In total, I identified 17 different contexts for language use from content analysis of my interviews (Table 3.3). I first calculated an amalgamated measure of general restriction for each speaker (Table 3.4), which can be interpreted as a rough indication of frequency of French usage, as well as an individual contextual restriction profile.

Table 3.3. Full list of contexts in which French is claimed to be used

(1) In childhood (father)	(10) At school (friends)
(2) Currently (father)	(11) At school (teachers)
(3) In childhood (mother)	(12) At work (working)
(4) Currently (mother)	(13) At work (in break room)
(5) In childhood (siblings)	(14) In the broader community
(6) Currently (siblings)	(15) Federal services (e.g. post office)
(7) Currently (spouse)	(16) Media consumption
(8) Currently (children)	(17) General preference
(9) Currently (friends)	

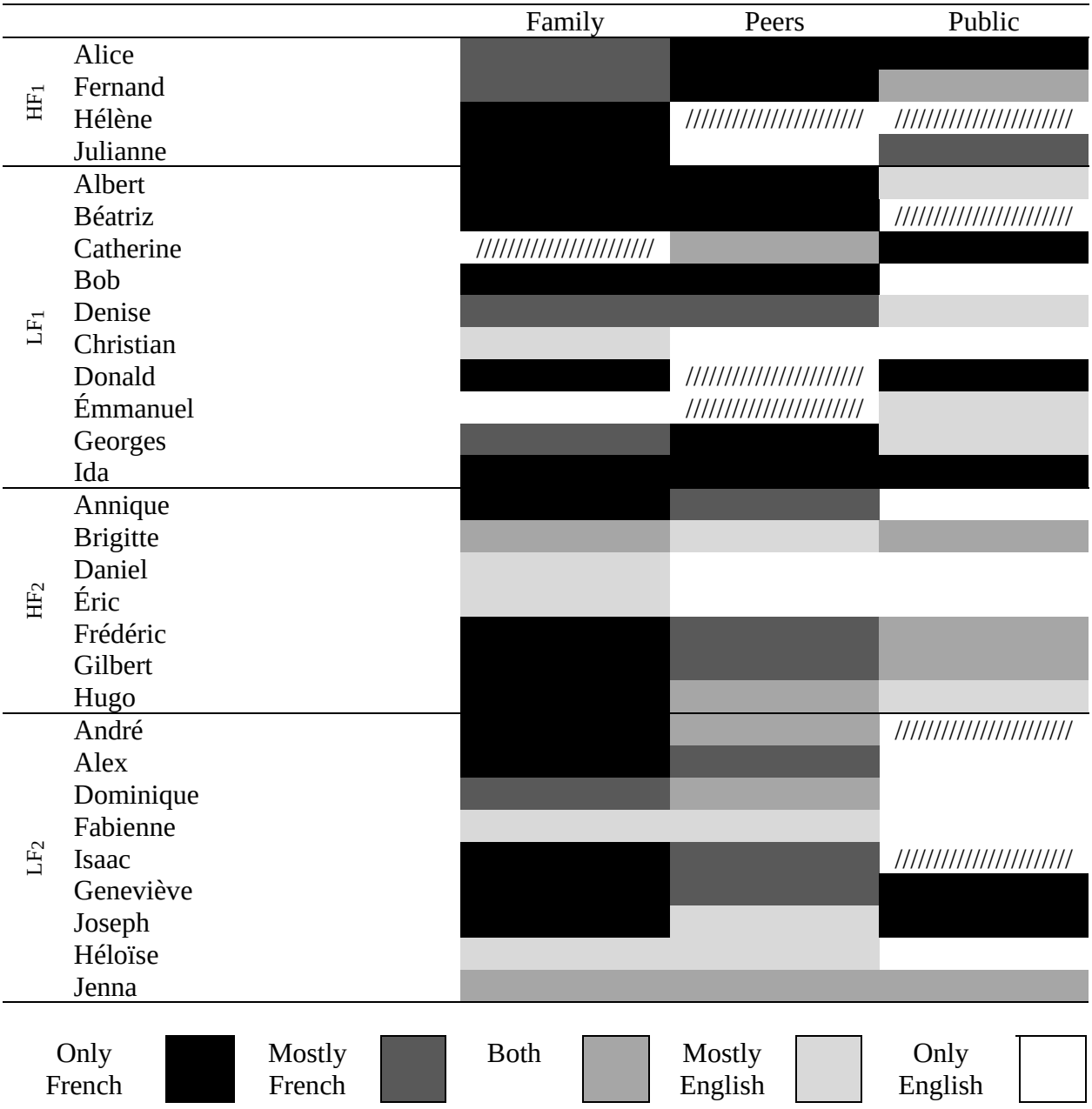
Table 3.4. Restriction index scores, by cohort

	HF_1		HF_2		LF_1		LF_2		TOT
OVERALL	13%	4	23%	7	33%	10	30%	9	30
Less restricted	21%	4	21%	4	32%	6	26%	5	19
More restricted	0%	0	27%	3	36%	4	36%	4	11

If contexts of use are generally important for linguistic outcomes (Mougeon and Beniak 1989; 1991; Mougeon and Nadasdi 1998; Sammons, Nadasdi, and Mougeon 2015), then it might be the case that some of the results can be explained by reference to specific contexts in which French is used (or not used). It also stands to reason that a broader range of contexts of French usage translates to more exposure to various usage norms that might exist within the community. Speakers whose use of French is largely confined to the home, for example, would

concomitantly be expected to receive less exposure to varieties of French unlike their own, and might then diverge from patterns attested for speakers who *are* in contact with other dialects. I compiled the individual restriction profiles by condensing the 17 identified contexts into three broader categories (Table 3.5): with immediate family (parents, children, spouse), with the peer group (siblings, friends), and out in the public domain (at work, when using services, in the broader community). Since language of media consumption is unlikely to cause accommodations (cf. Trudgill 1999, 6), it was not considered further.

Table 3.5. Language use in different contexts, by individual



Cross-hatching denotes categories that were not available for the individual in question.

Some more general observations about the contextual restriction of these speakers can be drawn from Table 3.5. There is a clear tendency to use both English and French with the peer group for the G_2 cohorts (perhaps self-evident from this group's level of bilingualism), whereas the G_1

cohorts tend either to use exclusively French or exclusively English with their friends. Some speakers have managed to maintain (near-) exclusive French usage in a wide range of contexts, including the public domain, whereas others seem to have found less occasion to use French. While it is true that, for some Victoria francophones, staying loyal to individual service providers who can communicate in French is of primary importance (even something of a point of pride), French in the public domain is probably not as common as this figure would lead us to believe. Because this information was culled from content analysis of the interviews, in some cases scores are based on responses in only one or two categories. Alice, for example, only provided information about her use of French at work. She happens to be employed in a francophone workplace (Brodeur), where she uses French exclusively. Table 3.5 should therefore be interpreted somewhat cautiously, as a general view onto contexts of language use across the different cohorts, rather than as a comprehensive assessment of the details of every speaker's pattern of language use.

It is common for G_2 speakers to undergo changes in contextual restriction once they leave school or leave home – the two bastions of French usage in Victoria. Daniel and Éric, Fabienne and Héloïse, and Dominique and Jenna (each pair a set of siblings) had all recently moved out of the family home at the time of their interviews, and the lower degree of French usage in those contexts reflects this shift. It is nonetheless important to underline the fact that, though some speakers are clearly contextually restricted – Fabienne, for example, reports using mostly or only English at home, with her peer group, and in the public domain – this does not appear to automatically result in appreciable effects on functional restriction, as (43) shows.

(43) Je me souviens comme une fois j'étais rentrée dans la salle de bain puis j'étais plus jeune puis elle était au secondaire. Puis j'ai poussé la porte puis elle avait pas fermé la porte à clé puis était juste là. Puis j'ai comme crié parce que j'av– [rire] *just like* plus parce que je savais pas qu'y avait quelqu'un puis j'étais comme « *oh my god* scuse ! » Puis elle avait comme un bonbon dans sa bouche puis elle m'a regardée puis était comme « *it's just a sucker*. » [VFC.LF$_2$.Fabienne.01/17/55]

'I remember like one time I walked into the bathroom and I was younger and she was in high school. And I opened the door and she hadn't locked it and she was just there. And I like screamed because I ha– [laughs] just like more because I didn't know that someone was there and I was like "oh my god sorry!" And she had like a candy in her mouth and she looked at me and was like "it's just a sucker."'

Fabienne generally expresses herself fluidly, without hesitation, and incorporates a number of phenomena associated with LF (use of the lexical item *salle de bain* to mean 'toilet facilities'[39]; *l*-deletion resulting in elision of the subject pronoun *elle* in *puis* [ø] *était juste là* and [ø] *était comme*, as well as in *qu'y avait*; shortening of the phrase *je m'excuse* to *scuse* [skyz]), as well as some more general features of colloquial French (pronunciation of 'puis' /pɥi/ as [pi]; deletion of the negation particle *ne* in *elle* [NEG-ø] *avait pas* or *je* [NEG-ø] *savais pas*). She is not alone in this either – the other speakers with low restriction scores are perfectly comfortable expressing themselves in French that seems quite natural. The quantitative analyses performed for this study will confirm the extent to which these superficial impressions reflect actual mastery of the underlying implicit patterns, but – at least on the surface – there is nothing to immediately

[39] In many European French varieties, *salle de bain* only designates the room equipped with a shower, bath, or sink for bathing, whereas the term *toilette* is reserved for the toilet facilities.

suggest that Victoria francophones are in late stages of attrition or have reduced functional linguistic capacity in French. Perhaps the contexts in which these speakers use *some* French – even if it is not the dominant language in these contexts – are sufficient to maintain the 'normal' level of French proficiency they would have acquired as a result of French being the dominant language in the home throughout childhood.

3.4.2. Contact with individuals and with infrastructure

To approach the question from a slightly different angle, I also devised two indices that quantify contact with individuals and contact with community infrastructure, using content analysis of the interviews for each speaker – a method reminiscent of the one used to compile the *Cumulative English Proficiency Index* of the Ottawa-Hull French Corpus (Poplack 1989). I attributed one point per speaker for each francophone institution (including francophone organizations, the school, the daycare, the parish, etc.) or recurring event, as well as each francophone individual or family that was named during an interview. These scores measure the extent of participants' involvement with community organizations, institutions, or events ('infrastructure' score), as well as the degree to which they interact with other individuals within the broader community ('individuals' score). I smoothed these raw scores by coding them as 'high' or 'low' according to their position relative to the mean of all scores: those higher than the mean are coded as 'high', while those lower than the mean are coded as 'low' (Table 3.6; Table 3.7; Table 3.8).

Table 3.6. Contact with infrastructure and with individuals, by individual

		Infrastructure	Individuals
HF_1	Alice	LOW	LOW
	Fernand	HIGH	LOW
	Hélène	HIGH	HIGH
	Julianne	HIGH	LOW
LF_1	Albert	HIGH	LOW
	Béatriz	HIGH	LOW
	Catherine	LOW	HIGH
	Bob	LOW	LOW
	Denise	LOW	LOW
	Christian	LOW	LOW
	Donald	LOW	LOW
	Émmanuel	LOW	LOW
	Georges	LOW	LOW
	Ida	HIGH	LOW
HF_2	Annique	LOW	LOW
	Brigitte	LOW	LOW
	Daniel	LOW	LOW
	Éric	LOW	HIGH
	Frédéric	HIGH	HIGH
	Gilbert	HIGH	LOW
	Hugo	HIGH	HIGH
LF_2	André	LOW	LOW
	Alex	LOW	LOW
	Dominique	LOW	HIGH
	Fabienne	HIGH	HIGH
	Isaac	LOW	HIGH
	Geneviève	LOW	HIGH
	Joseph	HIGH	HIGH
	Héloïse	LOW	HIGH
	Jenna	HIGH	LOW

Table 3.7. Contact with francophone infrastructure, by cohort

	HF_1		HF_2		LF_1		LF_2		TOT
OVERALL	13%	4	23%	7	33%	10	30%	9	30
High	25%	3	25%	3	25%	3	25%	3	12
Low	6%	1	22%	4	39%	7	33%	6	18

Table 3.8. Contact with francophone individuals, by cohort

	HF_1		HF_2		LF_1		LF_2		TOT
OVERALL	13%	4	23%	7	33%	10	30%	9	30
High	9%	1	27%	3	9%	1	55%	6	11
Low	16%	3	21%	4	47%	9	16%	3	19

Reviewing these scores reveals that different cohorts appear to engage with the community differently, with HF_1 speakers being the only group whose connections to the community largely tend to be via institutions rather than individuals, and LF_1 speakers being the only group to have primarily 'low' engagement according to both measures. This is perhaps indicative of the G_1 speakers having less of a 'built-in' local social network compared to G_2 speakers, who benefit more from the social opportunities provided by the school and who have a concomitantly higher degree of connections to individuals within the community. It also reflects how these speakers enter the community: G_1 speakers might be more likely to create a community connection by volunteering for an organisation or attending a special event; G_2 speakers have community connections through friend networks composed largely of other G_2 francophones.

3.4.3. Familiarity with school French

Given my earlier suggestion that explicitly taught rules of school French may have had a particular effect within the Victoria community, I also operationalized certain characteristics of the speaker sample that pertain to familiarity with or exposure to the rules of school French. All speakers in this sample have received roughly equal amounts of schooling in French, with a handful of exceptions. Twenty-seven of 30 speakers were educated exclusively in French-language institutions from kindergarten through to Grade 12, one G_1 speaker attended

anglophone high school in Montreal (Christian, LF_1), and two G_2 speakers (Brigitte and Daniel, both HF_2) switched to the French Immersion school system[40] after Grade 8. The speakers in my sample therefore can all be said to have received the same 'base' in school French instruction.

On top of this 'base', some G_1 participants likely have a high degree of familiarity with normative French via their profession. My sample contains an unusually high number of individuals in French-language education or otherwise language-centered professions (43%; N = 6/14), including teachers, translators, early-childhood educators, and educational assistants, which I have termed 'scholastic' professions. Individuals in scholastic professions not only speak French more often on a daily basis (as required for their job), but are also presumably familiar with the prescriptive rules of school French. Only two G_2 speakers are in language-centred professions themselves, but half (N = 8/16) have *parents* who are in language-centered professions, and might therefore be subject to a heightened awareness of school French rules via explicit correction of 'substandard' French by their parents. I therefore coded to reflect increased exposure to school French through parents and/or through profession (Table 3.9).

Incidentally, the high concentration of individuals in such language-centred professions is likely related to the fact that francophone workplaces that *aren't* somehow language-centred are exceedingly rare in Victoria. In fact, there are even a number of recruitment campaigns whose sole purpose is to hire teachers for Brodeur from Quebec and France. While this is a necessary move, in light of Victoria's small francophone population, to ensure that there is sufficient

[40] The reader should note that Brodeur does not follow French Immersion instruction methods, but rather the same L1 instruction style as would be used in any majority francophone context. Students are taught in French in all subjects (except English and other second-language classes) and are required to speak French when on school grounds.

faculty, this likely acts to inflate the number of established francophone individuals in Victoria who are in a scholastic profession, which in turn might heighten speakers' awareness of what is 'correct' and therefore possibly affect their linguistic choices – for instance, by avoiding any constructions which are known to be ungrammatical.

Table 3.9. Exposure to school French, by cohort

	HF_1		HF_2		LF_1		LF_2		TOT
OVERALL	13%	4	23%	7	33%	10	30%	9	30
High	20%	3	27%	4	20%	3	33%	5	15
Low	7%	1	47%	3	38%	7	27%	4	15

3.4.4. Other characteristics

There are certain other characteristics which were deemed hypothetically important to the linguistic behaviour of G_2 speakers. Among these is the ratio of HF-origin speakers to LF-origin speakers within the list of named individuals that was extracted for the 'individuals' score described above (Table 3.10).

Table 3.10. Proportion of HF-origin individuals in network, by cohort

	HF_1		HF_2		LF_1		LF_2		TOT
OVERALL	13%	4	23%	7	33%	10	30%	9	30
High	29%	4	29%	4	36%	5	7%	1	14
Low	0%	0	19%	3	31%	5	50%	8	16

Social class is one oft-cited factor mentioned in connection with the prestige or stigma of a given form: forms used by lower social classes tend to be stigmatised, whereas forms used by upper classes tend to be prestigious. As outlined in Section 1.2.3, previous research suggests a possible connection between prestige/stigma and the outcomes of dialect contact: speakers will avoid converging on stigmatised variants, and may overextend prestigious ones ('hypercorrection'). Standard classifications of social class or socioeconomic status often proceed on the basis of annual income or profession. However, there are no major differences within my sample of speakers along those lines – virtually all of them could be categorised as firmly middle class by this gauge of socioeconomic status. As such, I devised an alternative measure of this factor group based on the average price of a home within the neighbourhoods where participants reside (Table 3.11).

Table 3.11. Socioeconomic class (average house price in neighbourhood), by cohort

	HF_1		HF_2		LF_1		LF_2		TOT
OVERALL	13%	4	23%	7	33%	10	30%	9	30
High	59%	1	24%	4	47%	8	24%	4	17
Low	23%	3	23%	3	15%	2	38%	5	13

Finally, I hypothesise that duration of exposure to the Victoria context might have an effect on linguistic patterns, given the repeated and well-known findings that earlier age of arrival, longer duration of residence, or similar, increase the likelihood that speakers have converged (to the extent that they are able). In this study, I followed Otheguy, Zentella, and Livert (2007) in calculating the proportion of each speaker's life that would have been spent in Victoria, by

dividing years of residence in Victoria (self-reported during interviews) by each speaker's age (Table 3.12).

Table 3.12. Proportion of life spent in Victoria, by cohort

	HF_1		HF_2	LF_1		LF_2	TOT
OVERALL	29%	4		71%	10		14
High	33%	2		67%	4		6
Moderate	20%	1		80%	4		5
Low	33%	1		67%	2		3

In Chapter 2, I described a number of differences between the two generations of speakers of interest here, particularly as concerns the way in which they interact with the community. For example, some individuals maintain a network of francophone friends and participate in francophone extracurricular activities, which allows them to use French in a broader range of social contexts, whereas others confine their use of French to only one or two domains. These differences are operationalised using extralinguistic factors, but some of these must necessarily differ for each generation given the differences in their social realities. For instance, factors that capture degree of exposure to LF and HF, composition of the social network, use of French in the home, and so on, must be included for G_2 speakers as they are central to the research question at hand. Conversely, G_2 speakers do not differ according to proportion of life in Victoria, since

most G_2 speakers in this sample were either born in Victoria or arrived at a very young age[41] – though such a factor might prove to be explanatory of linguistic behaviour for G_1 speakers.

3.5. Summary

Ultimately, the goal of all this methodological machinery is to uncover whether G_2 speakers share norms, but – more particularly – to determine *which* norms they share, and whether they exhibit rates or rates as well as underlying patterns that reflect one (or more) of the models provided to them by the broader community. This is no easy task, given that all of the comparison cohorts share the same variants in most of the linguistic variables under investigation. Here is where the variationist approach and its comparative methods are so revealing as far as how the various putative sources of influence within this community (LF, HF, school French) might have affected the ultimate linguistic outcome at G_2. For example, it might be the case that one source of influence has been more important than the rest. If G_2 speakers largely tend to use implicit usage norms that 'match' those from only one benchmark (e.g. LF, as the majority dialect in the region, is a good candidate for this), it shows that communities where language use is extremely restricted are still able to develop into fully-fledged *speech* communities. On the other hand, we might find that G_2 speakers are not behaving as a unified or cohesive linguistic group: HF_2 speakers might very well align themselves with HF norms, while LF_2 speakers might instead align themselves with LF norms. A third possibility is that G_2 speakers might have turned to what they would have been explicitly taught in school (either from

[41] The two exceptions to this are Frédéric and Gilbert, who both arrived in Victoria at the threshold of adolescence (at age 11 and age 12, respectively). Since the general analytical procedure involves examination of linguistic behaviour at the individual level first, it should be obvious whether these two speakers pattern with the rest of their cohort or not.

prescriptive grammar, or from normative evaluations of some variants as being ‘better’ than others, or some overlap of the two). Either way, the *structure* of the variability will prove to be diagnostic, and this can only emerge from variationist analysis of diagnostic linguistic variables and internal and external conditioning .

Chapter 4. Assibilation as a window onto competing usage norms

The first variable I investigate is assibilation of coronal plosives. Assibilation was selected on the grounds of its comparability to a large number of previous studies in dialect contact which use categorical phonetic/phonological phenomena to determine whether G_2 speakers have aligned with one dialect over another. This variable can be considered effectively categorical in LF speech[42], as well as being one its more salient and emblematic phonological features (Walker 1984), but is virtually absent from HF speech. The G_2 sector of the Victoria francophone population was frequently and spontaneously identified by the participants in this study as having some kind of characteristic 'accent' that sets them apart from G_1 speakers. Although the subjectivity of these assessments is undeniable, and there is no mention of assibilation specifically as the characteristic that identifies G_2 speakers, it is interesting that many HF-origin participants identify G_2 speech as having a decidedly LF-like quality. The descriptors *québécois*, *Radio-Canada*, and even *Brodeur* – where the staff is predominantly LF-origin[43] – all suggest the presence or even prevalence of LF phonology. As such, this variable was chosen because its presence in G_2 speech would undoubtedly be remarkable to HF-origin speakers (particularly since this characteristic is associated with some stigma in their region of origin; see Section 4.2.2) and is immediately identifiable as an LF feature.

[42] Barring loanwords (e.g. *team*), which as a general rule do not have categorical assibilation.

[43] Though recall from note 25 that Brodeur also has a relatively high concentration of HF-origin individuals compared to the rest of the community.

(44) Elle prenait un accent québécois […] différent, donc, du mien.

[VFC.HF$_1$.Fernand.01/40/33]

'She was adopting a Quebecois accent […] different, then, from my own.'

(45) Ma mère elle dit toujours que– qu'on a l'accent 'Radio-Canada'.

[VFC.HF$_2$.Brigitte.01/10/35]

'My mum always says that we have a *Radio-Canada* accent.'

(46) On dit en blaguant que– qu'ils ont l'accent 'Brodeur'. [VFC.HF$_1$.Julianne.01/00/54]

'We jokingly say that– that they have the *Brodeur* accent.'

Furthermore, this feature is universal enough in LF varieties across Canada that any subtle fluctuations in assibilation production might be enough for LF-origin G_1 speakers to classify G_2 speech as somehow atypical or different from their own; for instance, as Donald does in (47), by explicitly labelling it as *franco-colombien*. Other community member assessments also seem to suggest that G_2 speakers deviate from LF expectations somehow (48)–(49).

(47) Oui, t'as un– t'as l'accent des franco-colombiens. [VFC.LF$_1$.Donald.01/09/40]

'Yes, you have– you have the *franco-colombien* accent.'

(48) [Ton français] a une– une couleur Colombie-Britannique. [VFC.LF$_1$.Beatriz.01/02/14]

'[Your French] has a– a British Columbia colour.'

(49) Leur parler québécois qui est différent de votre parler de jeune francophone qui a grandi ici. [VFC.LF$_1$.Albert.01/55/34]

'Their Quebecois speech that is different from your speech of a young francophone who grew up here.'

At any rate, given its status as a clear and emblematic marker of LF-origin phonology and its general phonetic salience, assibilation is an ideal feature to explore the question of whether G$_2$ phonology aligns with that of benchmark LF or benchmark HF, or even whether these speakers might have a distinct quality to their speech that distinguishes them from either of these models.

4.1. Background

Assibilation (as in (50)) is a form of lenition that results in the production of apicodental plosives with a secondary sibilant articulation when followed by a high front vowel (/y/, /i/) or glide (/j/, /ɥ/); for example, /ty/ '*tu*' is assibilated to [t^{s}y] in this environment.

(50) Crisse de tabarnak, elle m'a volé ma carte, la– l'**os<u>ti</u>e** [os**t^s**i] de machine. [VFC.LF$_1$.Donald.00/48/09]

'Goddammit, it stole my card, the– the fucking machine.'

Assibilation is fully phonologized in LF: apicodental plosives before high front vowels are categorically produced with the secondary sibilant articulation (Blondeau and Friesner 2014; Brasseur 2009; Côté 2014; Cox 1998; Detey 2010; Friesner 2010; Gess, Lyche, and Meisenburg 2012; Ostiguy and Tousignant 2008; Walker 1984), to the point of being "emblematic" of LF pronunciation (Berns 2013, 170). This contrasts markedly with the behaviour of these same

phonemes in HF, which are subject to more variability in the first place, and where the major variant (plosives *without* assibilation, as in (51)) is different from that of LF (Berns 2013; Corneau 2000; Junová 2018; Trimaille, Candea, and Lehka-Lemarchand 2012). The other two variants in the HF system are palatal affrication (52), which surfaces via the same process that yields assibilation[44] but yields a different phonetic outcome ([tʲ͡ʃ] instead of [tˢ])[45], and – *much* more rarely (Junová 2018) – assibilation itself (53).

(51) J'ai trouvé que c'était vraiment le même genre d'**architec<u>tu</u>re** [aʁʃitɛk**t<u>y</u>**ʁ].

[VFC.HF$_1$.Julianne.00/32/01]

'I found that it was really the same kind of architecture.'

(52) Il avait fait une **pein<u>tu</u>re** [pɑ̃**t<u>ʲ</u>**<u>y</u>ʁ] impressionniste. [VFC.HF$_1$.Julianne.00/40/22]

'He had done an impressionist painting.'

[44] The narrow stricture required to produce these vowels creates turbulence in the plosive release, which is hypothesized to trigger the presence of stridents following the plosive (Clements 1999; Hall and Hamann 2006; Kim 2001). The fact that the two surface variants (palatal affricates and plosives with secondary sibilant articulation) arise through the same (or similar) phonetic process is unimportant for the present study, since the focus is not on the underlying process, but rather on surface productions which clearly distinguish HF from LF.

[45] Some scholars group both phonetic outcomes into a single phenomenon, alternately termed either 'affrication' or 'assibilation'. I chose to distinguish them terminologically as well as methodologically in order to amplify the contrast between HF realizations (palatals, sibilants very rarely) and LF realizations (sibilants categorically).

(53) En fait ils ont acheté une **voi<u>t</u><u>u</u>re** [vwa<u>t^s</u><u>ʏ</u>ʁ]. [VFC.HF$_1$.Julianne.00/53/59]

'In fact they bought a car.'

Although there is some overlap between the implicit usage norms for HF and LF in that they both incorporate assibilation, the *rates* of this variant make it diagnostic of a difference between the two varieties nonetheless, since LF-origin speakers should produce assibilation categorically, whereas HF-origin speakers should opt for it sparingly or not at all. Conversely, HF-origin speakers should primarily use plosive variants – though palatal variants may also be used infrequently – while LF-origin speakers should categorically avoid plosive variants.

4.2. Benchmarks

4.2.1. Implicit usage norm in LF

Assibilation of coronal plosives is a well-established and virtually categorical phonological rule of LF. In addition to its regular occurrence within the bounds of the province of Quebec, the presence of assibilation in minority varieties of LF outside Quebec (e.g. in the Peace River region of Alberta [Walker and Canac-Marquis 2016], in Manitoba [Bérubé, Bernhardt, and Stemberger 2015], in Ontario [Durand 1993; Tennant 2012], etc.) suggests that this articulation is firmly entrenched for LF-origin varieties regardless of geographical distance from the Laurentian source. Given its persistence outside the Quebec hub, and the fact that Victoria is located in Canada, not France, it is reasonable to assume that speakers are receiving LF input from beyond the bounds of Victoria itself (e.g. through national francophone media), thereby increasing G_2 individuals' exposure to LF patterns. In addition, Victoria is home to a predominantly LF-origin

population that is continually renewed by new arrivals from Laurentian regions (cf. Chapter 2). It therefore seems likely that G_2 speech will turn out to be more LF-like than HF-like as far as assibilation, even despite the observations from some participants that G_2 speakers have an accent that is not typically '*québécois*' (cf. (47)–(49)). Of course, this is not to say that these observations are necessarily *wrong* – it might simply be the case that the G_2 deviation from LF expectations involves a phonological feature that is more subtle than assibilation (though obviously not so subtle as to preclude some HF-origin speakers from identifying LF-like characteristics in the speech of their children). The aim of this dissertation is not to identify each and every phonetic characteristic that might classify G_2 speech as being modeled on one of the three available options, but rather to probe areas where it will be easy to tell if they are aligned. Assibilation clearly represents one such area.

4.2.2. Implicit usage norm in HF

The plosive variant dominates in HF, with affrication as a minor variant, and assibilation as an even rarer option: Candea, Adda-Decker, and Lamel's (2013, 414) study of palatal versus plosive realization in a corpus of contemporary broadcast HF speech finds that palatal articulations occur at a rate of approximately 2-8%, with the sequence /tj/ (e.g. in *moitié, chrétien*) most likely to be palatalized, and /di/ (e.g. *vendredi*) the least likely. However, the presence of assibilation (or affrication) seems to be linked to particularities of oral cavity physiology or degree of vocal tract tension in the utterance of certain tokens. For example, Berns (2013, 241) provides a detailed acoustic comparison of the word *petit* produced by Canadian French and Metropolitan (Parisian) French speakers, which conclusively shows that production in the latter variety depends on the narrowness, tension, or shape of the vocal tract, whereas production in the former variety is phonologized and therefore categorical:

> "[…] in some cases, there is no genuine assibilation [in HF], only a certain amount of noise caused by the release of a tightly held obstruction as such. In yet other cases, real assibilation does arise, but it greatly differs in length (the Canadian assibilations may also differ in length, but contrary to speakers from France, the assibilated part is always clearly discernible, as if there were a kind of lower-bound, again demonstrating the categorical nature of the process in this variety)." Berns (2013, 241)

In other words, the implicit norm for HF strongly prefers plosive articulations, but does allow some lenition, which usually takes the form of palatal affrication except when vocal tract characteristics conspire to yield a 'tighter' assibilated pronunciation. Because the choice between these two options has more to do with vocal tract configuration and less to do with an implicit usage norm, selecting between assibilation and affrication is probably not dependent on internal linguistic conditioning per se. In contrast, selection of lenited variants versus plosive variants *is* conditioned by both internal and external factors, but as these are not diagnostic of a difference between HF and LF, I give them only a cursory treatment here. The *rate* of assibilation is the major diagnostic, but internal trends for HF speakers in Victoria may prove useful to confirm that their use of lenited variants is (or is not) in line with HF benchmark data.

Berns (2013) finds that voiceless /t/ is more likely than voiced /d/ to trigger some form of lenition (be it affrication or assibilation) – a finding that is corroborated by other studies (Candea, Adda-Decker, and Lamel 2013; Corneau 2000) – but this is the extent of the regular internal conditioning that has been uncovered. Other internal factors (e.g. following vocoid

[Corneau 2000], position of the segment within the word [Bento 2009; Junová 2016]) have turned up inconclusive or contradictory results in previous studies and are therefore not considered here. External factors such as age, gender, and socioeconomic class have also been associated with the selection of a lenited variant, with socioeconomic class frequently reappearing as an important factor. There seems to be a certain amount of stigma associated with use of palatalized variants in HF, which are alternately associated with "*la culture de rue*" in Paris (Jamin 2004), with "stigmatised young suburbanites" in Marseilles (translation mine; (Spini and Trimaille 2017), or with speakers of lower socioeconomic class more generally (Junová 2016). In these studies, however, this stigma does not appear to be against *assibilation* (or even lenition) specifically, but rather against the use of lenited variants by speakers who belong to a stigmatised group (e.g. Maghrebi-origin immigrants in Marseilles). That is, affrication or assibilation in a high-status speaker is unremarkable, but the same pronunciation in a low-status speaker is both salient and stigmatised (Blanchet 2016; Fagyal 2010; Vernet and Trimaille 2007). These results echo those from other fields, which suggest that normative associations of a speaker's accent are closely linked to evaluation of a speaker's social status: lower-status individuals are considered to have 'worse' accents than higher-status individuals – apparently even if the phenomenon in question is used by both groups (as is the case for assibilation in HF, cf. Blanchet [2016]). There is therefore unquestionably a stigma against the use of assibilation in HF – even if it is actually used by speakers from a range of socioeconomic backgrounds – which might block HF_1 speakers from adopting it, or possibly even cause them to pass on their avoidance of assibilation to G_2. In contrast to the situation in France, there is no stigma against assibilation for Canadians: all social classes, including the linguistic elite (e.g. *Radio-Canada* announcers; cf. Bigot [2008]; Cox [1998]; Reinke [2005]), categorically use

assibilation. The expectation, then, is that HF_1 speakers will not have adopted the LF pronunciation, thus providing G_2 speakers with an alternative to the LF option. It does seem unlikely, however, that G_2 speakers would acquire the negative associations with assibilation that their parents may have imported, since there are no such negative associations in Canada.

4.3. Variable context and exclusions

The variable expression of /t/ and /d/ with assibilation or affrication occurs immediately preceding the high front vowels or glides /i/, /y/, /j/, and /ɥ/, orthographically represented in four ways: <ti>, <tu>, <di>, and <du>. I performed an automatic search for these four orthographic strings within the ~200,000 transcribed words of my corpus (using ELAN v.4.9.4), which turned up over 6,600 tokens. To reduce the dataset to a more manageable size for manual analysis, I first filtered out all tokens that occurred within the first half-hour of the sociolinguistic interview to ensure that participants had 'settled in' and had accessed their most natural vernacular. According to Walker (1984, 91), *categorical* assibilation in LF only takes place within words and not across word boundaries; as such, I refined my token file to include the former contexts only. It might seem strange to purposely restrict my study of assibilation to an *in*variant context (given the *variable* framework employed throughout), but this was an intentional choice in order to maximize the possibility of contrast between HF and LF. While it would be in keeping with previous studies of these phenomena in HF to consider sibilant and palatal articulations together, I keep them separate for the initial analysis in order to determine whether they are treated as distinct options by the HF speakers in my sample.

Next, I limited the number of tokens for any given lexeme to a maximum of 10 per speaker and excluded all occurrences of '*tu-sais*' (which is a very frequent but fixed expression in LF[46] and therefore likely to skew results). Finally, I removed all accidental inclusions from the automatic extraction, such as English-origin material or transcription errors, and any poor quality or inaudible tokens.

Table 4.1. N tokens of assibilation in each cohort

	HF_1	LF_1	HF_2	LF_2	Total
N tokens	298	1105	465	583	2451

Since the articulations in question are fully distinguishable by ear[47], I relied on auditory measures to code the dependent variable, but I made sure to discard any tokens that were indeterminate on the basis of auditory information alone[48]. After these exclusions, the dataset contains 2451 tokens (Table 4.1).

[46] Murphy et al. (2016) report that native speakers of LF have strong intuitions that '*tsé*' is a single word rather than a reduction of two words as a result of fast speech.

[47] Spectral analysis of each token would be required to assess the specific phonetic properties of a segment (including whether the frication noise is from a glottal or oral source, cf. Berns [2013]), but the question addressed by my study is adequately answered using the coarser, audible distinctions between sibilant, palatal, and plosive realizations.

[48] A number of such tokens sound as though they fall into the category of *aspirated* plosives, i.e. the articulation that characterizes English plosives in certain phonological contexts (a finding also reported by Berns [2013]). They are excluded in this study because the auditory measure used to discern the variants is insufficiently reliable to distinguish aspiration from assibilation, but I make note of these as a possible avenue of future investigation since aspiration might be a salient phonetic feature that is identified as

4.4. Results

4.4.1. Overall

The distribution of each variant across the four cohorts reveals a clear split between LF-origin and HF-origin speakers (Table 4.2), but also between the two generations of the HF cohort, suggesting that the younger HF generation has collectively adopted the predominant LF pronunciation.

Table 4.2. Overall %assibilation for each cohort

	Sibilant		Palatal		Plosive		Total N
	%	N	%	N	%	N	
Overall	83%	2046	2%	43	15%	362	2451
HF_1	14%	43	**9%**	28	**76%**	227	298
LF_1	**99%**	1089	0%	0	1%	16	1105
HF_2	**76%**	355	3%	14	**21%**	96	465
LF_2	**96%**	559	0%	1	4%	23	583

The LF-origin speakers use the sibilant variant [t^s, d^z] almost exclusively, never use the palatal variants [$t^ʃ$, $d^ʒ$] (with one G_2 exception), and only make use of the plosive variant to a very minimal degree. As for the HF-origin speakers, the older generation strongly favours both the

'*franco-colombien*' by G_1 speakers. VOT has been shown in previous work (Robillard 2014) to participate in contact-induced change elsewhere in Canada, and may therefore reflect an influence from English – another possible explanation for why G_2 speakers might 'sound different' to G_1 ears.

plosive and palatal variants, whereas the younger generation makes use of both the plosive and palatal variants, but predominantly uses the sibilant variant.

The initial results therefore point to adoption of the predominant LF usage norm by all G_2 speakers. Although the HF_2 speakers do occasionally use the variants preferred by their parents (plosive, palatal), they largely use the variant preferred by their peers (sibilant). Given the invariance of the LF-origin cohorts, I do not elaborate further on the internal conditions which might govern their use of one variant over another – assibilation is clearly the preferred option – and proceed directly to an internal analysis of the more variable HF-origin cohorts.

4.4.2. HF_1 and HF_2 cohorts

Both HF-origin cohorts show far more variation within this paradigm than LF-origin cohorts, with each cohort using some of all three variants – which points up the *availability* of assibilation as an option for these speakers, even if they do not always avail themselves of it. The presence of variability is entirely as expected given benchmark data for this variable in HF, but there is a marked difference between the distribution of each variant in HF_1 and HF_2: the older generation uses predominantly plosives (76%; N = 227/298), whereas the younger generation uses predominantly sibilants (76%; N = 355/465). Clearly, HF_2 production is much closer to LF than to HF. Further supporting this assertion is the distribution of lenited variants (i.e., palatals and sibilants) in voiceless contexts versus voiced ones. Recall that previous research on HF suggested that voiceless plosives strongly favour lenited variants (Berns 2013, 232–33). Here, this effect is only seen for HF_1 (Table 4.3): HF_2 speakers use lenited variants at roughly equal rates regardless of plosive voicing (even slightly *dis*favouring lenition in voiceless contexts), whereas HF_1 speakers strongly prefer lenited variants with /t/ over /d/.

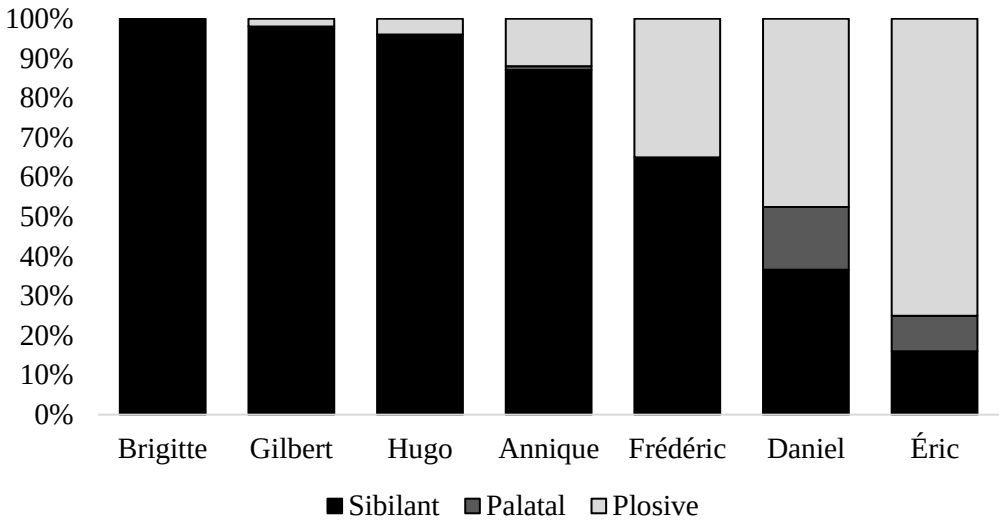
Table 4.3. Rate of lenited variants in voiceless (/t/) and voiced (/d/) contexts

% Lenition	/t/		/d/		Total N
	%	N	%	N	
Berns (2013)[a]	100%	83	24%	83	166
HF_1	**79%**	56	21%	15	71
HF_2	56%	208	44%	161	369
Overall	60%	264	40%	176	440

[a]: Rates and counts adapted from tables (154) and (155) in Berns (2013, 232–33), which include results for the sequences /ti/ and /di/ (i.e., no data is available for the sequences /ty/, /dy/, /tj/, /dj/, /tɥ/, and /dɥ/)

In other words, HF_1 speakers are exactly in line with benchmark HF trends, in that they strongly prefer plosives, but when they *do* select a lenited variant, it is very likely to be in a voiceless context. In contrast, the voicing effect is absent in HF_2 speakers and they opt for sibilants far more frequently, suggesting that their use of assibilation is approaching the phonologized, categorical behaviour of LF-origin speakers. What's more, it turns out that the overall rate of 76% assibilation for HF_2 speakers is concealing the fact that some individuals *do*, in fact, have phonologized assibilation: Brigitte, Gilbert, and Hugo's categorical (or very near-categorical) use of assibilation is perfectly in line with that of LF-origin speakers (Figure 4.1).

Figure 4.1. Distribution of each variant across individuals in the HF_2 cohort



4.4.3. Extralinguistic factors

Many speakers displayed virtually invariant (and therefore clearly phonologized) assibilation behaviour: all members of the LF-origin cohorts (except one, discussed below), and some members of the HF_2 cohort (Brigitte, Hugo, Gilbert) uniformly follow the LF community norm. Fisher's two-tailed exact tests confirm that generation and origin are both statistically significant to the use of assibilation (Table 4.4; Table 4.5), which demonstrates that this phenomenon is particular to LF-origin speakers regardless of generation, but also to G_2 speakers regardless of origin. In other words, these two social factors do a great deal to explain the tendency to select assibilation across the sample.

Table 4.4. Cross-tabulation (2x2 contingency table) of generation and assibilation variants

***p* < 0.0001**	**Sibilant**		**Plosive/Palatal[a]**		**Total N**
G_1	1132	81%	271	19%	1403
G_2	914	87%	134	13%	1048
Total	**2046**	**83%**	**405**	**17%**	**2451**

Significance calculated using Fisher's two-tailed exact test.

[a]: These variants are henceforth grouped together as 'HF' variants, a move justified by the fact that they pattern alike.

Table 4.5. Cross-tabulation (2x2 contingency table) of origins and assibilation variants

***p* < 0.0001**	**Sibilant**		**Plosive/Palatal**		**Total N**
LF	1648	98%	40	2%	1688
HF	398	52%	365	48%	763
Total	**2046**	**83%**	**405**	**17%**	**2451**

Significance calculated using Fisher's two-tailed exact test.

What might have impeded the phonologization of assibilation for the G_2 speakers whose use of this variant is not categorical (shown in Table 4.6)? Given the ubiquity and consistency of assibilation within the LF-origin speakers, we can assume that use of this variant is generally the norm within the broader Victoria community as well. As such, it seems reasonable to hypothesize that the speakers who only assibilate part of the time may do so because they have received insufficient exposure to the community usage norm (perhaps via lower levels of contact with individuals or infrastructure, via restriction, etc.); conversely, it may be the case that these

speakers have received higher exposure to HF than the speakers whose use of assibilation is categorical[49].

Table 4.6. Rates of each variant for non-categorical speakers

	Sibilant		Palatal		Plosive		Total
	%	N	%	N	%	N	N
Annique	87%	71	1%	1	12%	10	82
Frédéric	65%	46	0%	0	35%	25	71
Daniel	37%	23	16%	10	48%	30	63
Éric	16%	5	9%	3	75%	24	32
Jenna	77%	61	1%	1	22%	17	79

In order to test these possibilities empirically, I ran individual bivariate analyses for each of the extralinguistic factor groups against each of the three assibilation variants for the speakers in the HF_2 cohort, where much of the variability in assibilation selection was attested[50]. The reader will recall that multivariate analysis of these factors is not advisable for reasons of non-orthogonality adumbrated in Section 3.4, but bivariate analysis allows the analyst to determine whether any

[49] I exclude Frédéric from these deliberations, since at the time of the interview, he wore full orthodontic braces, which may have generally affected his pronunciation.

[50] The same procedure was applied to each of the cohorts, but the main analytical aim of this section is to determine whether extralinguistic factors might explain why some HF_2 speakers did not conform to the community norm of categorical assibilation. As such, analyses for the other cohorts are only adduced to confirm or refute tendencies that may emerge from the HF_2 data.

factors (in isolation) reach statistical significance, as well as the direction of effect that each of the factors might have on variant selection.

Although the data in Table 4.7 closely resembles a multivariate analysis in terms of presentation, the reader should note that each of the factor groups was run individually as a bivariate analysis to obtain *p* values. As such, no ranges or other calculations for magnitude of effect are computed, because these factor groups are not evaluated relative to each other and, as such, cannot be ranked. The percentages should be interpreted as indicators of direction of effect, with figures above the overall rate interpreted as a favouring effect, and those below the overall rate as a disfavouring effect. For instance, in the table below, the overall rate of the sibilant variant is 76%; it is favoured by lower exposure to school French (85%), but disfavoured by higher exposure to school French (63%), and this bivariate effect is found to be significant at the $p < 0.001$ level. It should also be noted that, where effects were categorical (either 0% or 100%), bivariate analysis in Goldvarb is not possible.

Table 4.7. Bivariate analysis of extralinguistic effects on assibilation selection for the HF_2 cohort

	Sibilant		**Palatal**		**Plosive**		**Tot N**
	76%	355	3%	14	21%	96	465
Exposure to school Fr.	***p* = 0.000**		*n.s.*		***p* = 0.000**		
Higher	63%	135	7%	14	30%	64	213
Lower	85%	220	0%	0	13%	32	252
Language with peers	***p* = 0.000**		***p* = 0.000**		***p* = 0.000**		
Both	87%	291	0.3%	1	13%	42	334
Mostly English	49%	64	10%	13	41%	54	131
Restriction	***p* = 0.000**		*n.s.*		***p* = 0.000**		
None	98%	42	0%	0	2%	1	43
Moderate	86%	249	0.3%	1	14%	41	291
Severe	49%	64	10%	13	41%	54	131
Individual contact	***p* = 0.000**		*n.s.*		***p* = 0.000**		
Higher	87%	220	0%	0	13%	32	252
Lower	63%	135	7%	14	30%	64	213
Infrastructure contact	***p* = 0.000**		*n.s.*		***p* = 0.000**		
Higher	97%	210	0%	0	3%	7	217
Lower	59%	145	6%	14	36%	89	248
%HF in network	***p* = 0.000**		*n.s.*		***p* = 0.000**		
Higher	63%	135	7%	14	30%	64	213
Lower	87%	220	0%	0	13%	32	252
Socioeconomic class	***p* = 0.006**		*n.s.*		***p* = 0.030**		
Higher	71%	129	7%	14	28%	50	213
Lower	80%	226	0%	0	16%	46	252

Light gray denotes significant favouring effects.

A number of interesting findings emerge from the bivariate results presented in Table 4.7. I focus here on those that can answer the hypotheses I emitted above about whether the HF_2 cohort's deviation from assibilation might be due to increased exposure to HF, or to decreased participation in the broader community. Table 4.7 shows that there is evidence for both effects. For instance, higher %HF in network correlates with a stronger tendency to use the plosive variant, and a categorical tendency to use the palatal variant – both of which are variants associated with HF speakers. As it turns out, this effect is also apparent in the LF_2 cohort (Table 4.8): higher %HF in network *strongly* favours use of the plosive variant.

Table 4.8. Bivariate analysis of effect of %HF in network on assibilation selection for the LF_2 cohort

	Sibilant		**Palatal**		**Plosive**		**Tot N**
	96%	559	0.2%	1	4%	23	583
%HF in network	***p* = 0.000**		***n.s.***		***p* = 0.000**		
Higher	77%	61	1%	1	22%	17	79
Lower	99%	498	0%	0	1%	6	504

Light gray denotes significant favouring effects.

Additionally, several the factors designed as alternate measures of participation in the community and/or usage of French also turned up as significant. Speakers who are unrestricted or only moderately restricted favour the sibilant variant (98% and 86%, respectively, compared to 76% overall), whereas those who are severely restricted strongly favour the plosive (41%, compared to 21% overall) and palatal (10%, compared to 0.2% overall) variants. This suggests that using French in a broader number of contexts in Victoria encourages HF_2 speakers to adopt the predominant, LF-oriented sibilant variant. Supporting this finding is the result for language with peers – those who mostly use English with peers disfavour the sibilant variant (49%) but

favour the plosive (41%) and palatal (10%) variants. Though it did not emerge as significant, the raw values associated with language in the home support my interpretation of the findings for language use with peers: those who currently use primarily English at home tend to use the plosive and palatal variants, whereas those who use French to some degree (equally with English, or predominantly) tend to use the sibilant variant. Reviewing the findings for other factors confirms this further: lower contact with individuals and infrastructure both have a significant disfavouring effect on use of the sibilant variant (63% and 59%, respectively), and a favouring one on use of the plosive variant (30% and 36%, respectively). Taken together, the general conclusion that can be drawn from these findings is that tendency to employ assibilation is highly sensitive to the degree to which an individual uses French, and the range of contexts in which they do so.

To illustrate these findings with concrete examples, let us briefly return to the non-categorical speakers shown in Table 4.6. Annique spent one of her childhood years living in France, and has always spoken French at home with her two HF-origin parents and brother, which may have been enough to cement plosives as an acceptable variant alongside sibilants – though the community norm has obviously contributed to her high rate of assibilation (87%; N = 71/82). This may be explained by the fact that she happens to be among those in the sample who maintain nearly exclusively French in her interactions within her peer group. On the other hand, Daniel and Éric – the only G_2 speakers to use more plosives than assibilation (cf. Table 4.6) – also spent time in France during their formative years, but unlike Annique, they maintain fewer ties with the francophone community, they no longer speak French with their parents (only one of whom is HF-origin), and they use exclusively English in interactions within their respective friend groups. Jenna's rates also support the conclusion that the friend group plays a

key role here, since she is the only G_2 speaker in this sample whose friend network contains a high ratio of HF-origin speakers: her slightly elevated rate of plosives relative to the rest of the LF_2 cohort reflects her exposure to HF in the friend group.

4.5. Summary and interim conclusions

Insofar as the overarching goal of this study is to determine whether G_2 speakers share implicit usage norms, then we have already reached a partial answer with this variable: most G_2 speakers *do* share phonological norms – which happen to be LF-oriented – with only a few exceptions. Overall, most speakers – including all of LF_1, nearly all of LF_2, and even some of HF_2 – conform to the predominant LF pattern of (near-)categorical assibilation. These results confirm the assertions by HF-origin participants that G_2 speakers' accents have a more LF quality, but do not bear out the suggestion by LF-origin participants that G_2 speakers somehow sound non-LF. As far as this highly salient and emblematic feature goes, G_2 speakers perform just like any other LF-origin speaker. For speakers who vary in their use of assibilation (three of the HF_2 speakers and one of the LF_2 speakers), I underlined a link between their exposure to HF and their current usage of French both with friends and family, and within the broader community.

What is especially noteworthy about these results pertains to *how* the G_2 speakers might have come to share implicit usage norms, and why a small number of speakers deviate from the general pattern. Firstly, since there is clearly a split at G_1 between HF- and LF-origin speakers, this demonstrates that the HF_2 speakers could not have acquired an LF-oriented norm directly from their HF-origin parents. They can only have acquired it through contact with LF speakers. And since G_2 speakers have a social network largely formed of other G_2 speakers, it must be the peer group – not parents or other G_1 speakers – who ultimately influence G_2 usage. Secondly, since there are a few HF_2 speakers who retain 'vestiges' of HF-oriented usage norms, in the sense that

they do not categorically assibilate, this is highly suggestive of the fact that the HF-oriented norms were acquired first, before gradually shifting towards the predominant LF-oriented norms of the broader peer group.

The primacy of the peer group in shaping norms can be illustrated by revisiting the case of Daniel and Éric, contrasted with that of Annique (the only three HF_2 speakers to have non-categorical assibilation) and Jenna. Annique has two HF-origin parents but continues to use French in her (largely LF-origin) peer group of other G_2 speakers; similarly, Jenna's peer group contains the highest ratio of HF-original individuals amongst the entire G_2 cohort. Daniel and Éric, in contrast, have one HF parent and one LF parent, but do not maintain a primarily francophone or even bilingual peer group. What this means is that, even though Daniel and Éric would have received LF-oriented input *directly* from their LF-origin parent (their father, Georges of the LF_1 cohort, who categorically assibilates), the fact that they do not use French regularly within their peer group (nor, indeed, with the broader community) seems to have precluded them from developing the categorical assibilation associated with LF-origin speakers. This must remain within the realm of conjecture as long as it only applies to a single variable, but it seems likely to be the case that similar trends are observable for the other variables.

Chapter 5. *Avoir/être*: A salient conflict between usage and the standard

In this chapter, we turn to the alternation between the auxiliaries *avoir* and *être*. This variable differs from assibilation in terms of level of linguistic structure (phonology versus morphosyntax), but also in terms of what it is designed to assess. Assibilation was included to test the relative influence of HF usage versus LF usage on G_2 speech, as these are the only two models of assibilation provided to G_2 speakers by the broader community (there is no prescriptive recommendation for or against assibilation, but it is still quite salient). The results for assibilation showed that, given the choice between HF and LF, G_2 speakers largely conform to LF. The *avoir/être* variable is not designed to test between HF and LF usage norms since – as I explain more fully below – the available evidence suggests that they are very similar. However, the (HF/LF) usage norm contrasts markedly with the prescriptive rule for this variable, and also prefers the avoidance of *avoir* in formal speech styles, which makes it the ideal diagnostic of whether G_2 speech is measurably influenced by school French, or whether it instead aligns with usage norms. This variable tests whether the cultural influence of the francophone school has translated into a linguistic influence as well – as suggested by participant evaluations of G_2 speech as evoking particularly 'standard' qualities (cf. (5), (27)–(29)).

Before proceeding with the rest of the chapter, it is necessary to clarify why LF and HF usage norms are not being treated separately for this variable. In contrast with assibilation, where LF and HF usage norms clearly differ, existing research does not specify any such dissimilarities for auxiliary selection. Much of the available literature speaks of auxiliary selection in LF only (Rea 2020; Sankoff and Thibault 1977; Willis 2000), which is largely conditioned by factors to do with the lexical identity of the verb: different verbs select *avoir* at different rates. However, what

little information exists about HF suggests that it does not pattern all that differently from LF. To my knowledge, this has not yet been assessed in a variationist framework (see also Rea [2020, 214]), but Kailuweit's (2011, 410) corpus study of HF finds that *avoir* occurs at least some of the time with the verbs *(re)passer* and *(re)monter*, though the exact rate of *avoir*-selection for each of these verbs is not provided. In addition, *Le bon usage*'s descriptions of *avoir* occurring with some verbs but not others (*partir*, *rentrer*, *rester*, *sortir*, and *tomber*) seems to align – at least on the surface – with the implicit norm for LF. It is clear that *Le bon usage* is written from the perspective of HF speakers (given their pains to identify where *non*-HF varieties might deviate from the 'norm'), and, as such, their descriptions of 'popular speech' almost certainly refer to HF. Since they observe that auxiliary selection varies for certain verbs in popular speech (see note 53), this strongly suggests that variability also exists in HF. Indeed, it would be very surprising if HF did not vary at all – though again, in the absence of published evidence, we cannot be sure of the extent nor the conditioning of this variability.

Since I cannot confirm on the basis of this alone whether the same verbs receive the *same* rate of *avoir*-selection in both HF and LF, I refrain from any strong claims of sameness between the two varieties. However, it stands to reason that the two may not be so different either – especially since we find virtually identical conditioning for this variable across a number of different North American varieties (in addition to the references for LF above, the same conditioning is found in Acadian French [Roussel 2016], Vermont French [Russo and Roberts 1999], and New England French [Stelling 2011]). In other words, the purpose of this variable is still served even if we cannot be entirely sure of its patterning in HF: it allows for a more direct comparison between school French and usage, since there is no discernible contrast between usage across so many varieties of French. If speakers still gravitate towards usage in this variable

– despite the relative simplicity/salience of the prescriptive rule for auxiliary selection – then it is strong confirmation of the primacy of usage over the standard.

5.1. Background

French combines either of the two auxiliaries *avoir* and *être*[51] with a verb participle to form different compound verb constructions, such as the *passé composé*. Although the vast majority of verbs take *avoir* categorically, including transitive and impersonal verbs, as well as the majority of intransitive verbs (Grevisse and Goosse 2016, 1125), *être* is used for the passive voice, for pronominal verbs, as well as for a small subset of intransitive verbs that express 'movement' or a 'change of state' (*ibid*., 1126), often colloquially referred to in pedagogical manuals with the mnemonic acronym 'Dr. and Mrs. P. Vandertramp' (termed 'VDT' verbs here). It is this last context that interests us here, since VDT verbs are the locus of robust variation in usage (54)–(55) despite very clear prescriptive mandate and normative condemnation against such vacillation in this context.

(54) Nous on **est**$_{[Ê]}$ **parti** quand on était en cinquième année. [VFC.HF$_2$.Hugo.00/05/39]

'We left when we were in fifth grade.'

(55) Parce que j'**ai**$_{[A]}$ **parti** juste après. [VFC.HF$_2$.Hugo.01/18/03]

'Because I left right after.'

[51] In this study I do not include what some grammarians call the 'semi-auxiliaries', e.g. *aller* (Grevisse and Goosse 2016, 1125).

Building on the seminal work of Sankoff and Thibault (1977), this variable has received considerable scholarly attention across a range of different sociolinguistic situations, including studies of French Immersion speakers (Knaus and Nadasdi 2001), of speakers with varying degrees of language restriction (Beniak and Mougeon 1989), of 'elite' speakers in formal settings (Bigot 2008; 2011; Villeneuve 2017), and of different regional Frenches including those spoken in Montreal (Rea 2016; 2020), Ontario (Willis 2000), Louisiana (Picone and Valdman 2005), New Brunswick (Péronnet 1991), Peninsular New Brunswick (Roussel 2016), Nova Scotia (Gesner 1978), Prince-Edward Island (King and Nadasdi 2001), Vermont (Russo and Roberts 1999), and France (Auger and Villeneuve 2017; Kailuweit 2011).

This variable seems to have attracted so much attention partly because it represents a clear instance where entire communities regularly contravene the prescriptive/normative recommendations in favour of the implicit usage norm[52]. Elements of the underlying conditioning of this variable are shared across a number of different varieties, except where selection of one variant or the other is effectively categorical (e.g. Prince Edward Island French, being essentially *in*variant in the tendency to use either *avoir* or *être*, cannot be subject to *variable* conditioning). Broadly speaking, selection of *avoir* is consistently favoured with some lexical verbs but not others (cf. Roussel 2016; Russo and Roberts 1999; Sankoff and Thibault 1977; Stelling 2011; Willis 2000). In addition, social factors that in some way operationalize an individual's social standing – be it in terms of amount of education (Rea 2016; 2020; Sankoff and Thibault 1977; Stelling 2011; Willis 2000), score on a linguistic marketplace index (Sankoff

[52] It was also a major contributor to the debate surrounding the *unaccusative hypothesis* in formal syntax (see Sorace [2000]). Variationist studies added a much-needed empirical element to contrast with the more abstract findings of prescriptive dictate and formal syntax.

and Thibault 1977), profession (Bigot 2011), or the more general 'socioeconomic status' (Knaus and Nadasdi 2001; Russo and Roberts 1999; Willis 2000) – also seem to point towards a strong *dis*inclination for the use of *avoir* in school French, since lower scores on each of these social-evaluative measures correspond to a higher rate of *avoir*. Conversely, higher social standing or more formal sociolinguistic situations incur a higher rate of standard or normatively-accepted forms (i.e., *être*; Bigot [2008, 271]).

Though the usage norms for this variable do not seem to differ much from study to study (even across typologically different varieties of French, such as Acadian and Laurentian), usage certainly clashes with the normative/prescriptive recommendations for this variable. Since most G_2 speakers in Victoria spend a significant portion of their French-speaking hours in the classroom (as opposed to in the home or on the playground), since the prescriptive rule for this variable is relatively straightforward and taught early in French classrooms (as early as grade 6; Ministère de l'Éducation de la Colombie-Britannique [2010]), and since the normative evaluation of *avoir* in these contexts is salient disapproval (cf. the external factors related to speaker status that were cited above), this is one area of school French that would theoretically be very easy for G_2 speakers to acquire. On the other hand, the implicit usage norm (described in Section 5.2.2) is both robust and extremely consistent across many studies (including studies of different varieties of French across North America and in Europe), which makes it a strong competitor against the prescriptive benchmark.

5.2. Benchmarks

5.2.1. Normative French

This variable is subject to a prescriptive rule and sociostylistic evaluations that happen to align as far as where and when *avoir* should be used. The prescriptive rule for this variable as it is outlined in *Le bon usage* (Grevisse and Goosse 2016, 1126) initially appears to be quite clear: use *avoir* for all compound verb constructions, except those in passive voice, those with pronominal verbs, and the intransitive verbs listed in (56), where *être* is compulsory (and consequently, *avoir* is prohibited). The most common of the 'mandatory-*être*' verbs are also included in the VDT set (underlined).

(56) *aller*, *arriver*, *décéder*, *devenir*, *échoir*, *entrer*, *mourir*, *naitre*, *partir*, *rester*, *retourner*, *sortir*, *tomber*, *venir*, *redevenir*, *rentrer*, *repartir*, *ressortir*, *retomber*, *revenir*, *parvenir*, *survenir*, *advenir*, *intervenir*, *provenir*

Note that although the verbs in (56) *must* take *être* when used intransitively per prescriptive rules, a number of them can be used transitively as well (e.g., *partir*, *sortir*, *tomber*...), in which case they categorically take *avoir*. For these verbs, *avoir* is *only* prescriptively acceptable when used transitively – *never* when used intransitively: *j'ai*$_{[A]}$ *parti*$_{[TR.]}$ *l'auto* is grammatically 'correct', but *j'ai*$_{[A]}$ *parti*$_{[INTR.]}$ *hier* is not[53].

[53] Though Grevisse and Goosse (2016, 1126) do make note of variability between *avoir* and *être* in popular speech for the intransitive verbs *partir*, *rentrer*, *rester*, *sortir*, and *tomber*.

However, there is another set of intransitive verbs where both *être* and *avoir* are allowed; with these, auxiliary selection does not necessarily depend on valency of the verb in question. Though some of the verbs in the latter set *may* also be used transitively (in which case *avoir* is categorical), they may also take *avoir* when used intransitively if the speaker's intent is to express an action rather than the resulting change of state (Grevisse and Goosse 2016, 1127). For example, while the transitive form *il a*$_{[A]}$ *changé*$_{[TR.]}$ *ses souliers* is perfectly acceptable, intransitive forms such as the ones in (57) are also both prescriptively accepted – except that the former would be prescribed for actions, whereas the latter would be prescribed for resultant states.

(57) il a$_{[A]}$ changé$_{[INTR.]}$ depuis la dernière fois qu'on s'est vu [= 'he **has** changed']

il est$_{[ɛ]}$ changé$_{[INTR.]}$ depuis la dernière fois qu'on s'est vu [= 'he **is** changed']

To sum up the prescriptive rules, *avoir* can be used in any compound verb construction – except with a small set of intransitive verbs that cannot be used transitively (e.g. *aller*, *arriver*, *venir*, etc.), which only take *être*. As for *être*, it can also be used with intransitive verbs that *can* be used transitively, but only if the intent is to express stativity.

In addition to the prescriptive rule, many studies have found a link between rate of *avoir*-selection and social standing, level of education, or formality of the speech context – regardless of how each of these external factors is operationalised (Bigot 2011; Rea 2016; Russo and Roberts 1999; Sankoff and Thibault 1977; Stelling 2011; Willis 2000). Recall from Section 5.1 that there is a clear association of *avoir*-selection with lower socioeconomic class, lower education, or less formal speech styles in studies of LF; in contrast, *être* is used more frequently

by speakers with higher status or in more formal speech styles (e.g. Bigot 2011; Sankoff and Thibault 1977; Stelling 2011; Willis 2000). Furthermore, Rea's (2016) longitudinal comparison of *avoir*-selection in Sankoff and Thibault's (1977) study to more contemporary Montreal French data reveals that *avoir* is on the decline – a finding which Rea attributes to gains in level of education across the broader Montreal community (Rea 2020, 240). This suggests that *avoir* usage may be especially sensitive to increased exposure to school French, and therefore that G_2 speakers will generally avoid it in my study.

5.2.2. Implicit usage norm(s)

Comparing usage with the rules of school French reveals some clear differences between the two. In particular, whereas prescriptive rules prohibit the use of *avoir* with intransitive verbs in the VDT set, implicit usage norms permit *avoir* with many of these verbs:

(58) Parce que j'**ai**$_{[A]}$ **parti** juste après. [VFC.HF_2.Hugo.01/18/03]

'Because I left right after.'

(59) Deux-mille-cinq, je pense c'est quand qu'on– quand on **a**$_{[A]}$ **sorti** [du campus Uplands]. [VFC.LF_2.Geneviève.00/26/26]

'2005, I think that's when we– when we left [Uplands campus].'

(60) Alors uh j– j'**ai**$_{[A]}$ **descendu** sur la grande route au lieu de prendre le métro. [VFC.LF_1.Georges.00/33/06]

'So uh I– I went down to the main road instead of taking the metro.'

This use of *avoir* where it is not permitted by prescriptive dictate is not random in usage: some verbs take *avoir* only rarely (e.g. *arriver*, *venir*), whereas others allow it far more frequently – and this finding has been replicated across many different studies of usage.

In addition, recall that while *être* must always be used with the list of intransitive verbs in (56), *Le bon usage* stipulates that it may also be *variably* used with another (much larger) set of intransitive verbs if the semantic intent of the speaker is to express a state (Grevisse and Goosse 2016, 1127). The same verbs should take *avoir* if the intent is to express an action. Per the prescriptive rule, the speaker of the two tokens below would presumably have 'intended' an active interpretation for (61), but a stative interpretation for (62).

(61) Après ça j'**ai**[A] **déménagé** en Ontario puis j'ai appris l'anglais là-bas peut-être. [VFC.LF$_2$.André.00/32/48]

'After that I moved to Ontario and I learned English there.'

(62) J'ai juste commencé à parler couramment [l'anglais] uh quand je **suis**[Ê] **déménagé** ici en troisième année. [VFC.LF$_2$.André.00/32/55]

'I just started talking [English] fluently uh when I moved here in third grade.'

While it is quite problematic to reconstruct speaker intent using spontaneous usage data, Sankoff and Thibault's (1977) work on auxiliaries in LF outlines some independent contextual cues that can indicate whether any given use of a verb is more action-like or more state-like – *without* relying on reconstructions of speaker intent. One such contextual cue, termed TEMPORAL INDICATORS OF COMPLETION, is used to distinguish action-like from state-like uses by

establishing whether the timepoint of reference of an utterance coincides with the time of the event itself ([+*complété*]), or whether the timepoint of reference is *after* the event has occurred or is ongoing ([-*complété*]). According to the prescriptive rule for this set of variable verbs, *avoir* should co-occur with [+*complété*] contextual cues, whereas *être* should co-occur with [-*complété*] contextual cues.

However, this one-to-one association of auxiliaries with stative or active readings does not account for the usage facts shown in many of the examples adduced above, including in (62) where the contextual indicators '*quand*' and '*en troisième année*' make a stative interpretation impossible. Take the pair of examples from the introduction to this chapter, (54) and (55), reproduced below:

(54) Nous on **est**$_{[\hat{E}]}$ **parti** quand on était en cinquième année. [VFC.HF$_2$.Hugo.00/05/39]

'We left when we were in fifth grade.'

(55) Parce que j'**ai**$_{[A]}$ **parti** juste après. [VFC.HF$_2$.Hugo.01/18/03]

'Because I left right after.'

Here we see that *être* has been selected with the [+*complété*] indicators '*quand*' and '*en troisième année*'. If prescriptive rules were operative here, *être* should denote a stative reading for this sentence, but the use of [+*complété*] indicators instead makes an active reading more felicitous (see also (62)). Similarly, there are a number of tokens where the use of *avoir* in the absence of disambiguating temporal indicators could be interpreted as either active *or* stative (63).

(63) Il a pas de moniteur de langue fait-que j'**ai**$_{[A]}$ **resté** en partie moniteur de langue, en partie lien avec la communauté. [VFC.LF$_2$.Joseph.00/20/16]

'There's no language monitor[54] so I stayed partly a language monitor, partly a community liaison.'

In other words, the prescriptive rule is certainly not hard-and-fast in usage – not only does *avoir* surface in usage where prescriptive rules prohibit it (i.e., with the intransitive verbs in (56)), there also appears to be no direct relationship between action/state readings and *avoir*/*être* usage, as suggested by Grevisse and Goose (2016). Instead, what most studies of auxiliary selection find (Rea 2016; 2020; Roussel 2016; Russo and Roberts 1999; Sankoff and Thibault 1977; Willis 2000) is that the use of *avoir* depends primarily on the verb involved: some verbs (e.g. *entrer, partir*, *rentrer*, *rester*, *sortir*, and *tomber*) select *avoir* far more frequently than others (*aller, arriver*, *mourir*, *naitre*, etc.). Since the verbs that generally favour *avoir* also happen to be those that have a parallel transitive use, and – conversely – those that generally disfavour *avoir* do not have a parallel transitive use, some scholars have emitted the hypothesis that it is this element of

[54] The term '*moniteur de langue*' denotes a specific role in many minority francophone schools as a cultural instructor or language assistant, whose duties may include planning and leading language-centred games or activities, or otherwise adding an element of francophone cultural development to support curricular instruction. This is a national Council of Ministers of Education, Canada (CMEC) program administered locally by the BC provincial government (see the program website at myodyssey.ca for more information). In many cases, these are individuals from outside the community who are stationed in a francophone school for one school year only, but Joseph – the speaker in this example – is a born-and-raised G$_2$ Victoria francophone.

transitivity that affects variant selection, rather than any independent lexical effect (e.g. Willis 2000). This possibility, as well as my operationalization of this factor group, is discussed in Section 5.4.

In addition to the usage norm for lexical verb, some studies suggest that *avoir*-selection depends on extralinguistic factors such as exposure to Quebec French, daily use of French, or amount of schooling in French (Stelling 2011). Exposure to Quebec French is a given for the speakers of my sample and cannot be tested directly, but the other two proposed factors are testable here. Stelling's study of New England Franco-Americans found much a lower likelihood of *avoir*-selection for speakers who had French as the sole medium of instruction. Since virtually all my speakers have the same amount of education in French (save three: two from HF_2 and one from LF_1), testing the effects of this factor group might not prove to be terribly revealing, but according to Stelling's findings, we should find that those speakers use more *avoir* than the rest of the sample. In addition, my speakers *do* vary in their daily use of French in different contexts, as shown in Section 3.4.1. Those who continue to use French on a daily basis (regardless of specific context, to align with Stelling's factor group) should use *avoir* less often than those who no longer use French as regularly.

5.3. Variable context and exclusions

The foremost concern for circumscribing the variable context is to isolate contexts where all the possible variants have referential equivalence – that is, where they all mean the same thing. As such, one necessary exclusion concerns syntactic configurations where VDT verbs are mandated to take *avoir*, such as in transitive uses (64) of VDT verbs. These were found to invariantly feature *avoir* in my data (N = 56) and as such were set aside prior to quantitative analysis.

(64) Je l'**ai**$_{[A]}$ **rentré** tout de suite ici parce que voulais pas qu'il soit dans mon garage.

[VFC.LF$_1$.Christian.01/25/13]

'I brought it in right away here because didn't want it to be in my garage.'

Other tokens were excluded because the auxiliary or the participle were not inflected (N = 57) or outright missing (N = 4), or because the utterance was incomplete or otherwise unclear (N = 16). A very small number were excluded because of the phonetic ambiguity between '*a*' and '*est*' (N = 4) in particularly fast speech, or because of difficulty in distinguishing auxiliaries in contexts of *liaison* (e.g. '*ils ont*/*ils sont*') (N = 1). I extracted all tokens of the different verbs in (56) that surfaced in my data (N = 17), as well as all tokens of derived forms of VDT verbs that were present in my corpus – *intervenir*, *redevenir*, and *repartir* – for a total of 19 lexical verbs under investigation.

Finally, I follow Sankoff and Thibault (1977) in excluding stative or copula contexts from my data, since *avoir* can never occur in those contexts. Their method for distinguishing copula/stative from auxiliary/active contexts is to establish whether the timepoint of reference of an utterance coincides with the time of the event itself ([+*complété*], as indicated by *le dimanche soir* and *à 2h du matin* in (65)), or whether the timepoint of reference is *after* the event had occurred or is ongoing ([-*complété*], as indicated by *toujours* in (66)).

(65) Le dimanche soir, uh à deux heures du matin, le *stroke* est– **est**$_{[ɛ]}$ **arrivé**.

[VFC.LF$_1$.Catherine.00/13/21]

'Sunday night, uh at two o'clock in the morning, the stroke happened.'

(66) Mais Hugo **est**[Ê] toujours comme **parti** quelque part *so* je le vois pas souvent. [VFC.LF$_2$.Héloïse.00/21/28]

'But Hugo is always like gone somewhere so I don't see him often.'

Circumstantial indicators of completion can also be used to make this determination. They were more straightforward to identify since most are simple adverbs or prepositional phrases. Each token was coded for presence of an indicator of manner (67), reason (68), cause (69), or goal (70).

(67) Il **a**[A] juste **passé** en saignant. [VFC.HF$_2$.Daniel.00/15/09]

'He just passed by bleeding.'

(68) Tout le monde **était**[Ê] **parti** parce que tous les camps sont six semaines. [VFC.LF$_2$.André.2.00/21/25]

'Everyone had left because all the camps are six weeks.'

(69) Mais je me souviens plus si on **était**[Ê] **sorti** à cause de ça. [VFC.HF$_2$.Hugo.00/04/25]

'But I don't recall anymore if we had gone out because of that.'

(70) Il **est**[Ê] **allé** à une autre ville pour acheter des oiseaux. [VFC.LF$_2$.Dominique.00/48/12]

'He went to another town to buy birds.'

Where the context did not permit a definite classification as [+/-*complété*], the token was coded as ambiguous. I never relied on the auxiliary itself to code a token as [+/-*complété*], which would have been circular. For instance, despite the fact that the auxiliary is *avoir* in the token in (71), which lends itself very easily to the assumption of a [+*complété*] reading, this token is coded as having a temporal indicator that marks *non*-completion (i.e., *toujours* 'always'; see also (66)).

(71) Sur l'île j'**ai**$_{[A]}$ toujours **resté** à Esquimalt ou proche. [VFC.LF$_1$.Albert.00/08/15]

'On the island I've always lived in Esquimalt or nearby.'

Thus, even though these sentences might appear 'ambiguous' as far as stativity, this factor group codes only for the presence (or absence) of indicators as outlined over several pages in Sankoff and Thibault (1977). Of the 1183 tokens in my dataset, only 25 appeared to be stative according to the coding protocol elaborated above, and only the token in (71) occurred with *avoir*. As such, it would seem that *avoir* is indeed strongly disfavoured with unambiguously stative contexts, so I opted to exclude these tokens from analysis, for a total of 1158 tokens.

5.4. Hypotheses about variant selection

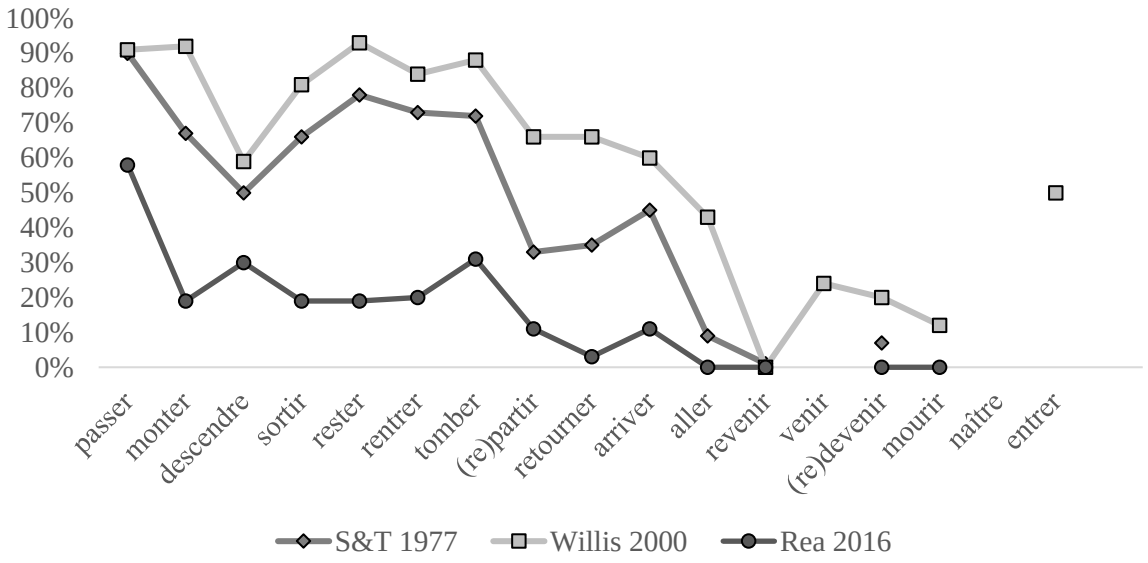
As I observed in the section on usage (Section 5.2.2), the primary diagnostic factor that distinguishes usage from prescriptive rules is co-occurrence of *avoir* with certain lexical verbs, which therefore forms the necessary focus of the present investigation. Little will be gained for my purposes by including tests of conditioning factors that are non-diagnostic. That is, rather than include some exhaustive collection of factor groups that have been previously tested for their effects on auxiliary selection in other studies, I made the analytical choice to only examine

those that can reveal if G_2 speakers are aligned with one of the three models available to them, and if so, which. In the case of auxiliary selection, only the factor group of lexical identity of the verb fits the bill. It should be noted, however, that pilot testing of a more exhaustive set of internal factors culled from previous variationist studies (including verb tense and mood, subject type, polarity, verb frequency, and presence of a direct or indirect object) revealed no major differences between the LF_1 cohort and the attested conditioning for benchmark LF.

In addition to testing lexical verb, I also examine whether there are any discernible effects for years of education in French or usage of French, which formed a part of the implicit norm in Stelling's (2011) study, as well as testing the effects of socioeconomic class in order to confirm whether they are also operative here. If speakers converge on the prescriptive rule, we should find low or no *avoir* use at all; if, in contrast, they use *avoir* at different rates according to different lexical verbs – and these lexical verbs correspond to those found to select *avoir* in benchmark studies of LF (e.g. Rea 2016; Sankoff and Thibault 1977)[55] – then this will constitute evidence that speakers have converged on the implicit usage norm. Some lexical verbs are consistently associated with very low rates of *avoir* (*venir*, *aller*...), whereas others are found to admit considerably more variation (*passer*, *sortir*, *rentrer*, *partir*...) – as can be seen from Figure 5.1.

[55] Of the studies discussed earlier, Sankoff and Thibault (1977) and Willis (2000) represent ideal comparison points with my G_1 data, and Rea (2016) with my G_2 data, since the other studies are of non-LF varieties (though North American nonetheless) or of specific speech modes (e.g. Radio-Canada speech; Bigot [2008]). It is clear from this figure that LF is currently undergoing a change in progress away from *avoir* in general (cf. Rea's results, from data collected in 2013, compared to those of Willis or Sankoff and Thibault, from data collected in the 1970s-80s, in Figure 5.1), but the pattern of variant selection is still quite discernible: some verbs select *avoir* far more frequently than others.

Figure 5.1. Rate of *avoir* according to lexical verb in benchmark LF data



Some studies test the different *attributes* of the lexical verb in isolation (as described in this section; these include frequency, parallel transitive use, parallel adjectival use), but it is also possible to simply test the rate of *avoir* for each lexical verb since the lexeme embodies each of the individual attributes. For example, the verb *aller* encodes high frequency, no parallel transitive use, and no parallel adjectival use. The effect of this verb on variant selection can therefore be assessed according to each of its individual attributes (high-frequency versus low-frequency; parallel transitive use versus no parallel transitive use; and so on), *or* it can be examined in direct comparison with other lexical verbs.

The effect of lexical verb frequency has been tested across several studies of auxiliary selection (Canale, Mougeon, and Bélanger 1978; Knaus and Nadasdi 2001; Roussel 2016; Russo and Roberts 1999; Sankoff and Thibault 1977; Stelling 2011; Willis 2000). Although very few

studies find statistically significant effects for this factor group[56], a general distributional trend is shared across many studies: the *low* frequency verbs are the ones that favour *avoir*. This is in line with Bybee's (2006) more general findings on frequency: irregular constructions that occur frequently in speech are more resistant to regularizing forces than those that only occur rarely. In this case, *être* is apparently resisting supplantation by *avoir* in only the most frequently-occurring lexical verbs, such as *aller*. The exception to this is in Willis' (2000) study: frequency was found to be not significant for her data, but her distributional results indicate that *être* is used more often in low-frequency verbs – the opposite of what the above studies found and an unexpected result according to Bybee's frequency generalization.

Most authors that include frequency as a factor group divide the lexical verbs into two groups – frequent and infrequent – but *how* they make that division varies from study to study. Russo and Roberts (1999, 75) use a published word frequency index for Parisian French to sort their verbs into each category (despite the fact that they were studying Vermont French, not Parisian French). Willis (2000, 31) uses frequency information directly recoverable from the fully-transcribed Ottawa-Hull French Corpus (Poplack 1989) from which she extracted her data. Knaus and Nadasdi's (2001, 295) frequency division relies on their dataset, but it pits a single verb (*aller*, by far the most frequent verb in their dataset) against *all* other lexical verbs. Once again, it is curious that such different methods of operationalizing frequency apparently lead to a uniform result across studies, especially since work on frequency (e.g. Walker 2012) shows that different ways of operationalizing frequency can lead to quite *different* results. This is a strong

[56] With the exception of Russo and Roberts (1999).

indication that it may not be the *frequency* of a given verb that matters so much as the use of the *lexical identity* of the verb itself.

In a similar vein, verbs that have parallel transitive uses have been found to be more likely to select for *avoir*. This effect is found in virtually every study of auxiliary selection, with the exception of Knaus and Nadasdi's (2001, 296) study of French Immersion speakers. A very similar effect – but for parallel *adjectival* uses – is also frequently linked to increased likelihood of *avoir* selection. Some studies, such as Roussel (2016), actually combine these two groups to identify verbs that can have *both* a transitive and an adjectival parallel use (versus, say, those that have neither). His study found that verbs where both uses are possible are by far the most likely to favour *avoir* (a finding also reported by Willis [2000]) but that verbs where neither are possible are *not* the *least* likely to favour *avoir* (verbs with only an adjectival use are the least likely). Once again, however, it looks as though the same effect could be obtained by just coding for lexical verb on its own, instead of coding for its component parts.

The reason for this is that the various categorisations seem less explanatory than the lexemes themselves, since verbs with the exact same properties show wildly different rates of *avoir* in previous studies. For instance, Willis (2000, 55) finds a rate of 88% *avoir* with the verb *tomber*, which has a parallel adjectival use, but only 24% *avoir* with *devenir*, which has the same feature. The verb *retourner* only gets 60% *avoir* in her study, but *passer, monter,* and *rester* – all other verbs which have both a parallel adjectival and parallel transitive use – get 91–94% *avoir*. Such examples could be multiplied. In my case, operationalizing these categories is further complicated by the fact that the verbs in my data were not orthogonal for this factor – 'adjectival use' and 'transitive use' both contain a single verb, whereas 'neither' contains two verbs and 'both' contains the rest. As such, I decided to include each lexical verb in isolation, rather than

test two or three factor groups that test frequency or parallel uses of the verb. These include the following lexical verbs: *aller*, *arriver*, *devenir*, *mourir*, *naitre*, *partir*, *passer*, *rester*, *retourner*, *sortir*, *tomber*, *venir*, *redevenir*, *rentrer*, *repartir*, *revenir*, *intervenir*.

I also include a test of whether the extralinguistic effects found in previous studies (that *avoir*-selection is strongly disfavoured in formal speech styles) are also operative here. I used topic in order to determine the speech style of an utterance, since certain topics such as education, religion, or metalinguistic commentary are deemed to elicit more formal or self-monitored speech, whereas childhood experiences, family traditions, or favourite pastimes produce less formal or vernacular speech. Nailing down the specific topic of an utterance is somewhat challenging, but the purpose of this exercise was ultimately to classify utterances as tending towards careful speech or tending towards casual speech. I identified 17 different broad topics, which are divided in Table 5.1 according to whether they were likely to instantiate more careful or more casual speech styles (e.g. personal or emotional narration versus soapbox speech or declaiming).

Table 5.1. Broad topics broached during sociolinguistic interviews

More casual	More careful
• Big decisions (moving to Victoria; making career choices; moving in with partner) • Mutual friends • Travel stories • Childhood experiences • Family stories, traditions • Scary experiences (bad injuries; getting ripped off; bear/cougar sightings) • Community involvement/community organizations/community members (incl. former teachers) • Hobbies/interests (music; cooking; games; TV; movies; books; clubbing) • Snow/cold weather experiences	• Controversial topics (homeless people; construction projects; climate change; researcher's own opinions; news; racism; politics; rental market; crime) • Education • Language (linguistic identity; linguistic prejudice; metalinguistic commentary) • Church/religion • Research project itself/interview • Career/work/work conflicts • History of the region • Polite offers (e.g. food/drink)

5.5. Results

5.5.1. Overall

A first surprising finding is that Victoria francophones *uniformly* prefer to use *être* over *avoir*: the overall rate of *avoir*-selection for all cohorts combined is only 4% (Table 5.2), with a low of 1% (HF_1) and a high of 6% (LF_2). These rates are very different from those attested for this variable in other studies, which range from 11-66% (Table 5.3), and are even lower than the rates attested in some studies of broadcast (i.e., more 'standard') speech (Villeneuve 2017).

Table 5.2. Overall %AVOIR for each cohort

	%AVOIR	N	Total
HF_1	1%	2	179
HF_2	3%	5	153
LF_1	4%	24	558
LF_2	6%	17	268
Overall	4%	48	1158

Table 5.3. Overall %AVOIR in previous studies

LF	Author	%AVOIR
Radio-Canada ('formal')	(Bigot 2008)	6%
Montreal	(Rea 2016)	11%
Radio-Canada ('informal')	(Villeneuve 2017)*	15%
Ontario (L2/Immersion)	(Knaus and Nadasdi 2001)	20%
Montreal	(Sankoff and Thibault 1977)	34%
Vermont	(Russo and Roberts 1999)	50%
Ontario-Hull	(Willis 2000)	66%
HF	Author	AVOIR
Vimeu	(Auger and Villeneuve 2017)	5%

*: Villeneuve's (2017) data is extracted from a Radio-Canada talk show that is said to instantiate 'more casual' broadcast speech. The rate of *avoir* reported in this table is adapted from data presented in Villeneuve's Table 2 (2017, 57).

5.5.2. Implicit usage norms for all cohorts

My first step was to check whether *avoir* is, in fact, used variably from verb to verb – as is the case in the implicit usage norms of the LF benchmark (cf. Figure 5.2; Rea 2016; Sankoff and Thibault 1977). As Table 5.4 shows, the verbs in my data that categorically avoid *avoir* closely resemble those that do the same in the implicit LF usage norm. For example, the verbs *mourir* and *naître* both categorically avoid *avoir* in most studies of LF, and I found this to be the case in my study as well. Other verbs, such as *devenir*, *redevenir*, *repartir*, *revenir*, *venir*, and *intervenir* are also (near-)categorical elsewhere; and in my study, they, too, all categorically avoid *avoir*. While the verbs that tend to avoid *avoir* are highly frequent in my data (e.g. *aller*, 29% of the data), removing these tokens (hypothetically) would actually change the overall rate of *avoir* very little. It cannot, therefore, be the case that the low rate of *avoir* in my data is attributable to a disproportion of verbs that (categorically) avoid *avoir*.

Table 5.4. %AVOIR and %POOL by lexical verb (all cohorts aggregated)

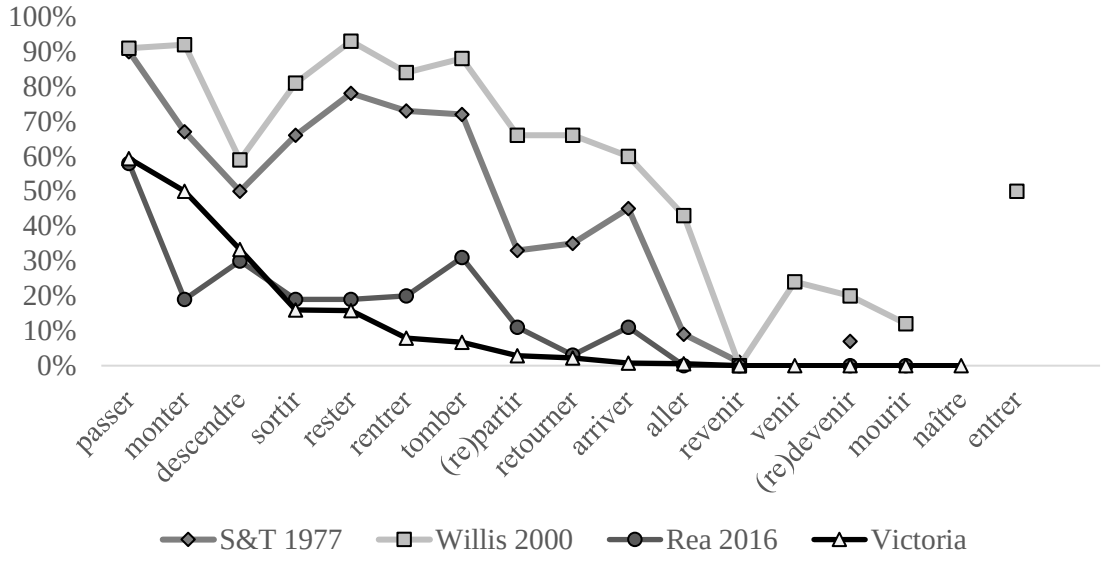
Verb	%AVOIR	%POOL*	N	Total
passer	59%	3%	19	32
monter	50%	>1%	3	6
descendre	33%	>1%	2	6
rester	16%	5%	9	57
sortir	16%	2%	4	25
rentrer	8%	3%	3	38
tomber	7%	1%	1	15
partir	3%	9%	3	107
retourner	2%	4%	1	45
arriver	1%	12%	1	142
aller	1%	29%	2	334
devenir	0%	4%	0	38
redevenir	0%	>1%	0	2
mourir	0%	1%	0	13
naître	0%	6%	0	70
repartir	0%	>1%	0	6
revenir	0%	5%	0	52
venir	0%	15%	0	168
intervenir	0%	>1%	0	2
Overall	4%		48	1158

*: %POOL is calculated on the basis of the full verbal pool, including categorical verbs, for each auxiliary. For example, *passer* represents 39% of the total *avoir* data, but only 1% of the total *être* data, and 3% of the data overall.

Very low rates of *avoir* are not unusual for these verbs (Russo and Roberts 1999; Sankoff and Thibault 1977; Willis 2000), with more recent studies finding that these verbs *never* co-occur with *avoir* (Rea 2016). Since the prescriptive rule mandates the use of *être* for these verbs, and the implicit usage norm also effectively only uses *être* for these verbs, this actually constitutes a *coincidence site*: lexical verbs that categorically take *être* are not diagnostic of whether speakers are converging on prescriptive rules or on the implicit usage norm. While the categorically *être*-selecting verbs cannot diagnose whether speakers are conforming to prescription or usage, the fact that other verbs variably select *avoir* in the Victoria data (though the overall rate is quite low) strongly suggests that these speakers have aligned themselves with LF usage. In other words, the categoricity of *aller*, *arriver*, *venir*, and so on *could* be attributable to prescriptive influence, but given the variable selection of *avoir* with at least some of the verbs, which is indicative of an influence of usage norms, the more likely explanation is that these verbs have 0% *avoir* because of usage norms, rather than because of speakers' awareness of (or desire to adhere to) the prescriptive rule.

Figure 5.2 plots the rate of *avoir* against each lexical verb included in this study (N = 17), for Sankoff and Thibault (1977), Rea (2016), and Willis (2000), as well as for my data. Observe from Table 5.3 and Figure 5.2 the decrease in overall rate of *avoir* from Sankoff and Thibault's (1977) study of Montreal French (34%), and Rea's (2016) study of the same variety (11%) – Rea's curve and the one for Victoria are very similar. Given that the rate of *avoir* has dramatically fallen in LF usage since Sankoff and Thibault's study some forty years ago, it might, then, be the case that Victoria francophones are simply in line with this trend away from *avoir* usage.

Figure 5.2. Rates of benchmark *avoir* usage in benchmark LF data and in Victoria data



As Figure 5.2 shows, the rate of *avoir* depends on the lexical verb in the LF benchmark data, and the Victoria data is well-aligned with this trend as well – despite having much lower rates overall. Rea's data – being the most contemporary – is probably the best comparison point for my own data. Although our rates of *avoir* for each lexical verb are not identical, there are a minimum of 'rerankings' in the hierarchy of verbs that permit *avoir*, or the apparent 'rerankings' can actually be attributed to a lack of tokens in the Victoria data. For instance, Rea's data shows that *tomber* is still among the verbs that selects *avoir* the most frequently (second only to *passer* in the Montreal data), but my data has 'reranked' *tomber* to a lower position – which appears to be a legitimate difference between the hierarchies in our two datasets. However, the apparent reranking of *monter* (which ranks low in Rea's but high in my data) is almost certainly due to the very small token count for *monter* in my data (N = 6; 50% *avoir*), which has the effect of artificially inflating its rate.

Examining whether this effect of lexical verb holds across each cohort shows that there is some fluctuation in *avoir* according to lexical verb within each cohort, but that this is clearly a feature that is more associated with the LF-origin cohorts (Table 5.5). Where the HF-origin cohorts generally avoid *avoir* (though HF_1 does so more than HF_2), the LF-origin cohorts use *avoir* far more frequently across a range of different lexical verbs: variability in HF_1 is confined to *passer* and *retourner*, whereas the LF_1 cohort uses *avoir* in every verb except one (*retourner*). This suggests that the implicit usage norm is not uniform at G_1, but we cannot forget that the benchmark usage patterns according to verb in HF are not explicitly available from any studies, and what little information there is suggests that it is very similar to LF. I therefore hesitate to read too much into these rate differences, which may also (at least partly) be due to the fact that HF_1 speakers have a third of the tokens that LF_1 speakers do.

Table 5.5. Selection of *avoir* according to lexical verb, by cohort

Lexical verb	Overall		HF_1		HF_2		LF_1		LF_2	
%AVOIR	4%	48/1158	1%	2/179	3%	5/141	4%	24/558	6%	17/268
passer	59%	19/32	33%	1/3	43%	3/7	50%	7/14	80%	8/10
monter	50%	3/6	—	—	—	—	40%	2/5	100%	1/1
descendre	33%	2/6	0%	0/3	—	—	100%	1/1	50%	1/2
rester	16%	9/57	0%	0/8	0%	0/7	26%	6/23	16%	3/19
sortir	16%	4/25	0%	0/1	20%	1/5	18%	2/11	14%	1/7
rentrer	8%	3/38	0%	0/3	0%	0/6	13%	1/8	10%	2/21
tomber	7%	1/15	—	—	—	—	10%	1/10	0%	0/5
partir	3%	3/107	0%	0/12	6%	1/17	2%	1/53	4%	1/24
retourner	2%	1/45	8%	1/13	0%	0/3	0%	0/15	0%	0/13
arriver	1%	1/142	0%	0/12	0%	0/15	1%	1/100	0%	0/16
aller	1%	2/334	0%	0/52	0%	0/59	1%	2/142	0%	0/81

As for the G_2 cohorts, if we compare the rates-by-verb for each of the two G_2 cohorts with the benchmark LF usage data from Sankoff and Thibault (1977) and Rea (2016), we see that the (slightly) increased rate of *avoir* usage in G_2 cohorts relative to G_1 cohorts appears to be due to the propensity of some verbs (e.g. *passer*) to have increased their tendency to be *avoir*-selecting compared to G_1. That is, the G_2 rate of *avoir* is not higher than G_1's because they have applied *avoir* to a wider range of lexical verbs, but rather because they use more *avoir* in some of the verbs 'supplied' to them as *avoir*-selecting by the LF usage norm. I return to this finding in the interim discussion section below.

Another internal factor that it is possible to test (though not in a multivariate framework, given the very low token counts) is speech style. Although this has been found elsewhere to be a significant contributor to variant selection (Bigot 2008; Sankoff and Thibault 1977; Stelling 2011; Willis 2000) it seems extremely unlikely that it could be entirely responsible for the very low rates of *avoir* seen here. Because sociolinguistic interviews are generally informal in tone – and mine are especially so since I have pre-existing relationships with many of the participants – the predominant use of *être* in this data is not likely to be entirely due to a preponderance of formal topics. The results indicate that style has no effect whatsoever on variant selection for both the LF-origin cohorts, but that HF-origin cohorts categorically avoid *avoir* when using a careful speech style (Table 5.6).

Table 5.6. Effect of speech style on selection of *avoir*

	HF_1		HF_2		LF_1		LF_2	
Overall	2%	2/107	4%	5/119	6%	24/382	9%	17/199
More casual	5%	2/94	5%	5/108	5%	14/258	9%	16/187
More careful	0%	0/13	0%	0/11	8%	10/124	8%	1/12

This suggests that the HF-origin cohorts might have some negative association with *avoir* usage that the LF-origin cohorts lack. It also shows how G_2 speakers are not perfectly uniform, since HF_2 speakers avoid *avoir* in careful style, but LF_2 speakers do not.

5.5.3. Extralinguistic factors

Fisher's two-tailed exact tests (Table 5.7; Table 5.8) show that generation is not significant, but that origin is. This constrasts with assibilation, where both were statistically significant. Previous

studies are mute as to whether HF and LF pattern identically for this variable (though the available evidence strongly suggests that this is the case), but here we see the first suggestion that *avoir* usage may vary according to origins.

Table 5.7. Cross-tabulation (2x2 contingency table) of generation and *avoir/être* variants

***p* = 0.1702**	***avoir***		***être***		**Total N**
G_1	26	4%	711	96%	737
G_2	22	5%	399	95%	421
Total	**48**	**4%**	**1110**	**96%**	**1158**

Significance calculated using Fisher's two-tailed exact test.

Table 5.8. Cross-tabulation (2x2 contingency table) of origins and *avoir/être* variants

***p* = 0.0328**	***avoir***		***être***		**Total N**
LF	41	5%	785	95%	826
HF	7	2%	325	98%	332
Total	**48**	**4%**	**1110**	**96%**	**1158**

Significance calculated using Fisher's two-tailed exact test.

However, though there are significant differences between the *rate* of *avoir*-selection for HF- and LF-origin speakers, this should not be interpreted as confirmation of the fact that underlying patterns of *avoir* selection are different for speakers of different origins (indeed, where there is sufficient data to ascertain patterns, they appear quite similar) – just that HF speakers and LF speakers opt to use it at different rates, regardless of generation. Why do HF-origin speakers not use *avoir*-selection to the same degree as LF-origin speakers (especially HF_2 speakers, who should

presumably be especially sensitive to trends predominant in the broader community)? The effect of speech style seen in Table 5.6 might provide a plausible explanation (i.e., HF-origin speakers regardless of generation associate *avoir* with stigma, whereas LF-origin speakers have no such negative associations), but confirming this first requires a closer examination of the other extralinguistic factors (Table 5.9).

As it turns out, no factors (or but a single factor) emerge as significant for all cohorts except LF_2; the HF_1 and HF_2 cohorts barely use *avoir* at all, and therefore have very little in the way of conditioning, and the factor weights for the LF_1 cohort all hover around .50, indicating that extralinguistic factors matter very little to variant selection for this particular cohort. As such, the discussion of extralinguistic effects focuses on the results for the LF_2 cohort only.

Table 5.9. Bivariate analysis of extralinguistic effects on *avoir* selection for the LF_2 cohort

	avoir		**Total N**
	6%	17	268
Exposure to school French	***p* = 0.017**		
Higher	3%	5	154
Lower	11%	12	114
Language with peers	***p* = 0.000**		
Mostly French	17%	12	71
Both	1%	1	94
Mostly English	4%	4	103
Restriction	***p* = 0.000**		
None	17%	12	71
Moderate	3%	4	119
Severe	1%	1	78
Individual contact	***p* = 0.005**		
Higher	9%	16	174
Lower	1%	1	94
Infrastructure contact	***p* = 0.006**		
Higher	16%	9	58
Lower	4%	8	210
Socioeconomic class	***p* = 0.000**		
Higher	13%	15	113
Lower	1%	2	155

Light gray denotes significant favouring effects.

This table reveals a very similar story to the one uncovered for assibilation: use of *avoir* seems largely dependent on the degree to which LF_2 individuals connect with the community. For instance, individuals who are unrestricted favour the use of *avoir*, as well as those who have more contact with individuals and with infrastructure, and those who use mostly French with peers. This suggests that, like assibilation, use of *avoir* is contingent on exposure to this implicit usage norm from the broader community. Although there was no effect of speech style on LF_2 *avoir* usage, the extralinguistic results suggest that there may be a relationship between *avoir* usage and exposure to school French: those who have greater familiarity with school French are significantly more likely to avoid *avoir*.

Other effects are somewhat more subtle and reflect complex interactions between the social factors in this community. For example, at first glance, it may seem counterintuitive that elevated socioeconomic class should favour *avoir* usage (especially if this variable is also affected by factors like exposure to school French and speech style), but this turns out to be explained by an interaction with language at home and contact with individuals in the community: most LF_2 individuals with higher socioeconomic class use exclusively or predominantly French in the home, and all have a high degree of contact with individuals in the community. Consequently, it seems clear that these individuals do not use more *avoir* because they belong to the higher socioeconomic class, but rather because they have increased exposure to informal usage via their continued use of French in a number of different contexts.

5.6. Summary and interim conclusions

To summarize the findings for *avoir/être*, LF-origin speakers of both generations clearly retain the same major conditioning factor as was found to be important in noncontact LF, since the rate of *avoir* fluctuates according to lexical verb, and the same lexical verbs favour *avoir* in my data

and elsewhere. The HF-origin speakers are somewhat divided: HF_1 speakers do not follow the same lexical rule that is in place for LF (but also only have two tokens of *avoir*), whereas HF_2 speakers seem to abide by it at least partially. They certainly do not contravene it. The LF_2 speakers seem to have adopted the LF usage norm, except they may have overextended it slightly, since they permit *avoir* in virtually all of the same verbs but actually use higher rates of *avoir* than would be expected from benchmark LF in some of the verbs (particularly *passer*). The HF_2 speakers, while not perfectly in line with the LF_2 cohort, are also not inconsistent with it. They, too, use *avoir* in a range of verbs (though this range is slightly more restricted than in LF_1 or LF_2, but certainly less restricted than HF_1), and also use comparatively more *avoir* where it is licensed by the LF usage norm (e.g. in *passer*).

This variable was included as a test of whether implicit usage norms or prescriptive rules play a more important role in G_2 norms. Although all speakers minimize *avoir* overall (which would certainly suggest an awareness of the prescriptive rule, though see the discussion below about the change in progress for LF), the external factors for LF_2 provided evidence that greater participation in the community leads to a higher likelihood of using *avoir*, which suggests that vernacular *avoir* usage is available primarily to those who have a higher degree of exposure to the broader community. In addition, I uncovered an effect of speech style (for HF-origin speakers only) which disfavoured *avoir* in careful speech, as well as an effect of exposure to school French (for LF_2 only) whereby speakers with increased familiarity with school French used less *avoir* – both of which seem to indicate that some stigma against *avoir* usage might exist within different sectors of this community.

What might explain the very low rate of *avoir* seen for Victoria speakers across the board? A number of possibilities present themselves, but it is interesting that – despite very little exposure

to the implicit usage norms of LF (given the lack of *avoir* at G_1) – there is evidence to suggest that it has been noticed and adopted at G_2 with only minor differences. I had proposed above that this might be the continuation of a change in progress away from *avoir* that has been attested for Montreal (Rea 2016), which I find no evidence to rule out here. An alternative explanation is that there may not have been that much *avoir* in this community to begin with: recall that my sample is, generally speaking, highly educated and many G_1 individuals in my sample and in the broader community are employed in professions that require good command of prescriptive rules. It might be the case that these speakers were already well-versed in eschewing *avoir* when they arrived in Victoria, but that where they *do* choose to use it reflects the implicit LF usage norm.

Another explanation is that any uses of *avoir* in their environment are highly salient as 'red flags' to G_2 ears – especially given the contexts in which the broader G_2 cohort most often hears and uses French (i.e., the school), where *avoir* would not be explicitly taught, and likely even labelled as a mistake and explicitly corrected. My G_2 sample is exposed to French both at school and in the home, and the contrast between these two models might have encouraged G_2 speakers to somewhat overgeneralize the use of *avoir* relative to the G_1 usage norms. We saw from the distribution of *avoir* across different lexical verbs that G_2 speakers have not expanded the list of verbs with which it is 'acceptable' to use *avoir*; instead, they appear to have increased the rate of *avoir* with some verbs (most notably *passer*) but not others – presumably under the implicit belief that *avoir* must be in some sense 'acceptable' with *passer*. The selection of this particular verb cannot be an accidental choice, either, since *passer* belongs to the set of verbs that is, in fact, licensed to take *avoir* (under specific conditions) by both the prescriptive enterprise *and* the LF usage norm, and it is even one of the verbs explicitly mentioned by Kailuweit (2011) as admitting variability with *avoir* in HF. It therefore does not overtly violate any of the available

models, and might thus be perceived as the ideal candidate for *avoir* usage by G_2 speakers – especially in the absence of 'normal' levels of exposure to usage, which might 'correct' this generalization by the G_2 cohort.

As far as the research objectives of this study, it is clear that G_2 speakers in both cohorts are aligned as far as rates (though whether these can be unequivocally attributed to school or prescriptive influence is not clear, given Rea's recent findings of a general LF trend away from *avoir*), and also clearly show some familiarity with the implicit usage norm – thus confirming their linguistic cohesion for this variable. The results for *avoir/être* also provide a slight contrast to those of assibilation, which in turn opens up a new angle of investigation, since assibilation clearly showed that G_2 speakers align themselves near-categorically with LF usage patterns. Here, we see that they definitely incorporate elements of LF patterns (as evidenced by their variable use of *avoir* with verbs that license this auxiliary in benchmark LF), but it also seems plausible – given the very low rate of *avoir* – that prescriptive teachings have mitigated the tendency to use *avoir* with the same wide range of verbs as is attested for benchmark LF communities. Because this variable is subject to a change in progress which happens to be aligning the usage norm more closely with the prescriptive rule – in that both mandate the use of *être*, not *avoir*, with these verbs – it is not possible to disentangle these two possibilities. However, the other variables may prove to be more conclusive in this respect.

Chapter 6. Future temporal reference: A three-way conflict with non-salient variability

The first two analysis chapters each described variables where G_2 speakers have a choice between only two models. Assibilation focused on the competing influences of HF versus LF usage norms, but not prescriptive rules (because no such rules exist for assibilation), whereas *avoir/être* tested the competing influences of usage norms (which characterise both LF and HF) versus prescriptive rules. The findings from these two variables, taken together, suggest that LF usage takes precedence over HF usage, and *avoir*'s salience was proposed as a possible contributor to the attested effects of speech style and of exposure to school French – both of which suggest that use of *avoir* in contexts where it is proscribed has some negative associations within this community.

The variable of future temporal reference is a test of whether this general finding still holds when G_2 speakers must choose from all three models simultaneously. In addition, the first two variables were highly salient in different ways (assibilation because of its emblematic association with LF, and *avoir/être* because of overt instruction from the school that categorically mandates the use of one variant but prohibits the other). This is not true of future temporal reference, where both the prescriptive rules and the usage norms are significantly less salient. Since salience is known to contribute to social indexation, and if negative associations cause speakers to avoid a given variant (as was the case for *avoir/être*), then it is probable that the extralinguistic factors of interest in this study will have weaker or no effects on a less salient variable like future temporal reference. As a corollary of this, it might be the case that salience (and associated social indexation) is what allows G_2 speakers to overcome the general lack of opportunities for speaker

interaction and converge on a shared norm. Since future temporal reference operates largely below the level of conscious awareness (both in terms of its prescriptive rule, which we will see is rather nebulous, and in terms of its usage norms), we might find that G_2 speakers do not converge in the same way that they did for the first two variables. However, it is also true of future temporal reference that its variability in usage is largely conditioned by one major (and several more minor) factors, making it an excellent test case for the degree to which G_2 speakers are sensitive to the underlying conditioning of variability.

6.1. Background

To refer to future time, French has three primary constructions from which to choose, two of which were sufficiently frequent to include in this study[57]: the PERIPHRASTIC FUTURE (PF; (72)) and the SYNTHETIC FUTURE (SF; also sometimes called the inflected future; (73)).

(72) Est-ce que après ça ils **vont aller**$_{[PF]}$ en Afrique? [VFC.LF$_1$.Albert.01/02/19]

'After that will they go to Africa?'

[57] Recent work on HF by Gudmestad et al. (2020) suggests that the third option used to express future time in French, the *present tense future* (P), occurs far more frequently in HF than it does in LF (compare their finding of 26% present tense future [N = 87/311; Table 1] to the rate found in LF by Poplack and Dion [7% FP; N = 242/3559; Table 7]). This variant was highly infrequent in my data – as is the case for many other LF communities – and was consequently set aside. While the relative rarity of FP in my data might be considered a first indication that Victoria francophones, on the whole, are more in line with LF than with HF, it may also be simply due to the preponderance of LF-origin speakers in this sample.

(73) Si je suis pour aller à messe, j'**irai**$_{[SF]}$ pas en anglais. [VFC.LF$_1$.Albert.00/26/03]

'If I'm going to go to mass, I won't go in English.'

As is the case for each of the variables in this study, the prescriptive rule for the use of these variants is the same for both HF and LF, as the monocentric ideols of French grammar generally do not permit different sets of rules for different varieties of French. However, the implicit usage norms for each of HF and LF differ both from the prescriptive rule and from each other. These differences are described in more detail in the sections that follow.

6.2. Benchmarks

6.2.1. Normative French

Recall from Chapter 3 that a major goal of the prescriptive enterprise is to achieve form-function symmetry in their recommendations; that is, they seek to factor out the inherent variability that is present in naturalistic speech by associating one form to one function, and/or a separate function to each form. Despite the apparent simplicity of the future temporal reference paradigm, grammarians have not been generally successful in their attempts to attribute the use of these competing verbal forms to specific functions (cf. Poplack and Dion 2009). The difficulty of ascribing one-to-one functions to these forms likely stems from the fact that the functions of future temporal reference forms have evolved over time in natural, spontaneous speech – as language structures are wont to do. These shifts in function are not well-handled by the prescriptive enterprise, who tend to view language as immutable. Comeau and Villeneuve (2016, 232) outline the trajectory of these forms from Vulgar Latin, which reveals that the periphrastic

future was initially used to denote spatial movement rather than temporal movement (Fleischman 1982), though at present it has fully grammaticalized into a marker of futurity.

Though neither of the variants appears to be explicitly prohibited by prescriptive grammars, several agree that the synthetic future has a 'neutral' quality (cf. Poplack and Dion 2009, 568), thereby suggesting that the periphrastic future must be somehow marked or perhaps more specific in function. One older grammar (amongst many others) suggests that the periphrastic future be restricted to "*la langue familière*" ('popular speech'; Baylon and Fabre [1973, 126]), which seems to be the prevailing opinion in more contemporary works as well. For instance, the recommendation for future temporal reference provided by *Le bon usage* positions the synthetic future as the default or neutral form (which they simply term *le futur*), and the periphrastic future as an alternative generally employed in speech:

> "Le futur est fortement concurrencé, surtout dans la langue parlée, soit par le présent, soit par des périphrases au moyen de semi-auxiliaires : *aller* principalement, mais aussi *devoir*, *vouloir*, etc." (Grevisse and Goosse 2016, 1193)
>
> 'The future is strongly contested, especially in speech, by the present tense, or by periphrases with semi-auxiliaries: primarily *aller* [go], but also *devoir* [must], *vouloir* [want], etc.'

In a thorough investigation of prescriptive dictates for future temporal reference over five centuries (1500–1999), Poplack and Dion (2009) identified 20 different prescribed functions assigned to the synthetic future and 19 to the periphrastic future, but found most of these to be

largely ephemeral, non-contrastive with the functions for the other forms, or even contradictory in some cases (*ibid.*, p. 564–8). Furthermore, prescriptive recommendations appear to have multiplied in recent decades, with 80% of prescribed functions for one or the other variant appearing in grammars published after 1950 (*ibid.*, p. 565). Within this quagmire of rules, only one emerges as relatively consistent: many grammarians over the last few centuries have agreed that the periphrastic future should be used for proximate future events (Poplack and Dion 2009, 564). In other words, according to the prescriptive rule, the periphrastic future is more strongly associated with events that will occur in the near future, which – per the doctrine of form-function symmetry – would leave the synthetic future as the form associated with the distant future. As it turns out, Poplack and Dion (2009) found that the operationalization of this prescriptive rule in multivariate modelling was only weakly significant, which suggests that the temporal distance rule is not actually what conditions future temporal reference variability in natural LF speech.

6.2.2. Implicit usage norm in LF

The implicit LF usage norm for future temporal reference has completely diverged from normative injunction. Though it does not align with the rules of prescriptive grammar, it has been shown to be very stable in LF communities across time (Poplack and Dion 2009; Sankoff and Wagner 2020; Wagner and Sankoff 2011). Of primary importance in the LF usage norm is the effect of polarity on variant selection: the synthetic future is restricted almost categorically to negative polarity utterances, whereas the periphrastic future is used in both negative and affirmative contexts. Given the number of studies of future temporal reference in LF across many different speech communities (Blondeau 2006; Grimm 2010, 2015; Poplack and Dion 2009; Poplack and Turpin 1999; Wagner and Sankoff 2011), it is unsurprising that linguistic factors

other than polarity have been found to be significant as well[58], but the *characteristic* constraint in the LF grammar for future temporal reference is that of polarity. In addition to polarity, one other factor appears to be diagnostic of a difference between HF and LF: adverbial specification. Though the effect of this factor group is marginal at best in LF (Poplack and Dion 2009, 573), it appears to have far greater importance in HF (Gudmestad et al., [2020, 180]; see also Villeneuve and Comeau [2016] for Vimeu French), which justifies its inclusion here. In studies of LF, this factor group is operationalised in one of two ways: binary presence versus absence of a temporal adverb (e.g., Poplack and Dion 2009), or ternary specificity of the adverb, non-specificity of the adverb, and no adverb (e.g., Blondeau and Lemée 2020). For greater comparability with studies of HF, I opt for the latter coding protocol (see Section 6.4.3).

There is evidence that the synthetic future has implicit positive associations, which is demonstrated in studies that assess future temporal reference in more formal varieties or contexts, such as the one by Blondeau and Labeau (2016). In their dataset of scripted weather reports for broadcast news, they find a far higher rate of synthetic future than in studies of informal LF. Although prescriptive grammars identify the synthetic future as being 'neutral', it would seem that this is interpreted in the implicit usage norm as 'prestigious' – although since Blondeau and Labeau were studying recitation of prepared texts, this could also be an effect of written versus oral sources. There are, however, a number of other indications that the synthetic future is associated with prestige somehow; for instance, Poplack and Dion (2009, 577) find an effect of speech style (pronouns) whereby the synthetic future is favoured in formal speech

[58] In addition to the polarity constraint at the top of the ranking, other significant effects include temporal distance, speech style, and adverbial specification (Poplack and Dion 2009), as well as subject type and contingency (Grimm 2010; Wagner and Sankoff 2011).

styles. In a similar vein, Wagner and Sankoff (2011) find an effect of subject type that favours the synthetic future with the formal pronoun '*vous*'. Some studies also report that the synthetic future was favoured by individuals with higher socioeconomic status or education (Villeneuve and Comeau [2016]; Wagner and Sankoff [2011][59]; though see Poplack and Dion 2009, who do not find such an effect).

There were indications that *avoir/être* is sensitive to effects of speech style and familiarity with school French, which act to depress the rate of *avoir* relative to benchmark data. However, the low rate of *avoir* could also be explained by the trend in progress away from *avoir* that is underway elsewhere in Canada. Both are plausible explanations for the low rate of *avoir*. Fortunately, it is possible to disambiguate between these two explanations with future temporal reference. The synthetic future is evaluated as 'neutral' – so not downright prestigious, but less marked than the periphrastic future – but there is also a "vigorous" trend in progress for future temporal reference whereby use of the synthetic future is rapidly losing ground to the periphrastic future (Poplack and Dion 2009, 572). In other words, if my sample were tailoring their linguistic behaviour to use higher rates of whatever variant is more prestigious (or at least, less stigmatised), we would expect to see *higher* rates of the synthetic future than those in benchmark data, but if they are instead following changes in progress away from certain variants that are underway elsewhere in Canada, we would expect to see *lower* rates of the synthetic future than those in benchmark data[60]. Thus, this variable can shed some light on the question of whether Victoria speakers may be aligning themselves with changes in progress elsewhere in the

[59] In the latter study, this effect was confined to older speakers only.

[60] Since the benchmark data that is available predates my corpus by many years, it follows that the rate of synthetic future would be *lower* than (and not *on par* with) benchmark data.

LF-speaking world (fairly likely, given the LF alignment exhibited for assibilation and *avoir/être*) – to be interpreted as an alignment with the implicit LF usage norm – or whether they instead adapt their linguistic behaviour to incorporate higher rates of 'prestigious' forms.

6.2.3. Implicit usage norm in HF

Variationist (or even quantitative) studies of future temporal reference in HF are rather thin on the ground compared to those for LF, but the findings that *are* available indicate that the synthetic future is selected far more frequently in HF than it is in LF (e.g. compare Roberts' [2012] rate of 41% synthetic future in HF to Poplack and Dion's [2009] 20% in LF). Nevertheless, there are some surface similarities between the implicit usage norms in HF and LF. In particular, the polarity constraint plays a synthetic future-favouring role in both varieties (Fleury and Branca-Rosoff 2010; Gudmestad et al. 2020; Roberts 2012), but this finding is not as robust in HF as it is in LF. For the HF system, Gudmestad et al. (2020) actually found polarity to be the *least* important factor in their analysis, with adverbial specification instead emerging at the top of the ranking. This contrasts very starkly with the implicit norm for LF. In LF, polarity is the most important constraint, with adverbial specification only having a very minor effect; in HF, these effects are reversed, and polarity has a lesser effect, but adverbial specification emerges with the highest magnitude of effect. Roberts (2012) found that polarity was the *only* conditioning factor in this variety – after testing for a number of internal factors, including adverbial specification. However, Roberts' results still represent a contrast with LF, in that the polarity constraint in HF is by no means categorical.

Like LF, there also seems to be a connection between prestige and use of the synthetic future in HF. For example, subject type has a similar effect in HF as it does in LF: formal address pronouns favour the synthetic future (Roberts 2012). In addition, Roberts' study also finds that

speakers with a higher level of educational attainment (university degrees) use significantly more synthetic future than those with less education. These results (though not significant in Roberts' study) seem to suggest that the implicit norm in HF might include similar associations of prestige with use of the synthetic future as was found for LF, which makes it non-diagnostic of differences between the two. Since it can nonetheless contribute to the discussion around whether sociostylistic evaluations shape linguistic norms for Victoria francophones, it is included in the analysis below.

To summarize, the major difference between the implicit norms for HF and for LF is that polarity emerges as the only significant factor in HF, but actually has a much weaker effect than it does in LF. Either it is ranked below another factor (as in Gudmestad et al. [2020]), or its effect is nowhere near categorical (as in Roberts [2012]). In contrast, the LF system is governed primarily by polarity, but also by a number of other internal factors.

6.3. Variable context and exclusions

The variable context for future temporal reference must be circumscribed to include only those contexts that express *future time* – in Poplack and Turpin's (1999, 143) words "any and all reference to a state or event occurring posterior to speech time" – since it is only in this context where the two variant forms are equivalent in referential function (see also Zimmer [1994]). That is, forms that feature future morphology but express something other than future time (as well as those that do not admit all the variants) must be excluded, such as habituals[61] (74), imperatives

[61] Including *gnomics*, i.e. forms that refer to "general truths or tendencies" (Sankoff and Evans Wagner 2006, 280).

(75), forms that indicate spatial movement (76), frozen forms or expressions (77), or protases of hypothetical *si*-complexes (78).

(74) Une fin de semaine par mois ils **vont faire**$_{[PF]}$ un voyage à l'extérieur.
[VFC.LF$_2$.Fabienne.00/01/43]
'One weekend a month they take a trip away.'

(75) Tu peux dire à ton chien « **va avertir**$_{[PF]}$ uh– uh– la maison qu'on rentre pas ce soir. »
[VFC.LF$_1$.Catherine.00/31/48]
'You can tell your dog "go tell uh– uh– the house that we're not coming home tonight."'

(76) Fait-que je m'en **vais donner**$_{[PF]}$ un examen là-bas. [VFC.LF$_1$.Catherine.00/13/16]
'So I'm headed to give an exam over there.'

(77) Québécois, le peuple de chialeux, on **va dire**$_{[PF]}$. [VFC.LF$_1$.Albert.00/19/52]
'Quebecois, the nation of whiners, let's say.'

(78) Si tu **vas écouter**$_{[PF]}$ l'entrevue de ta grand-mère. [VFC.LF$_1$.Albert.02/13/51]
'If you go listen to your grandmother's interview.'

Unlike Comeau (2015), Roberts (2012), and Grimm (2015), I opted not to exclude tokens of reported speech (79)–(80). Even though temporal information pertaining to proximity of the event in time is not recoverable from these tokens, they were nonetheless variable in their

selection of synthetic future and periphrastic future in my corpus. I did, however, exclude quotations which are clearly direct echoes of another's speech or of writing, such as in (81).

(79) Puis je dis : « hop je pense ça **marchera**[SF] pas. » [VFC.LF1.Catherine.00/45/08]

'So then I say "Oh I think it won't work."'

(80) Ouais les gens m'avaient dit ça, ils ont dit : « **vont** te **casser**[PF] ton antenne ! » [VFC.LF1.Catherine.01/38/30]

'Yeah people told me that, they said "they'll break your antenna!"'

(81) « *Dès qu'on **arrivera***[SF], » hein ? « *Dès qu'on arri–* » elle– elle écrivait « *des cons* » c-o-n-s, tu-sais ? [VFC.HF1.Hélène.01/03/31]

' "As soon as we arrive," eh? "As soon as we arri–" she– she would write "*des cons*" ('idiots') c-o-n-s, you know?'

Once these exclusions are taken into account, this dataset contains 733 tokens.

6.4. Hypotheses about variant selection

The factor groups described in this section have each been hypothesized to affect variant selection individually, but the crucial differences between the implicit usage norms for future temporal reference in HF and LF – as well as the differences between the implicit usage norm and the prescriptive rule – are only observable from the relative ordering of magnitude for factor groups within a multivariate analysis, the direction of effect for the factors in each group, and to a lesser extent from the rates of each variant. In other words, the factor groups themselves are not

diagnostic in isolation; it is their position relative to each other that is important for future temporal reference. For instance, if the temporal distance constraint tops the ranking for any of the cohorts in this analysis, we would conclude that the cohort in question is aligned with the prescriptive usage rules, but not with HF or LF usage norms, as these do not rank temporal distance above other constraints. In a similar vein, an HF-aligned ranking would have a weak effect of polarity, whereas an LF-aligned ranking would have a strong or near-categorical effect of polarity.

6.4.1. Polarity

By far the strongest implicit constraint in LF is that of sentential polarity: the synthetic future variant is categorically preferred in negative polarity contexts (82).

(82) Ah il **va parler**$_{[PF]}$ le français, il **aura**$_{[SF]}$ pas le choix. [VFC.LF$_1$.Catherine.01/50/46]

'Ah he'll speak French, he won't have a choice.'

This effect has been found in several variationist studies of LF (e.g., Emirkanian and Sankoff 1985; Grimm 2010; Sankoff and Evans Wagner 2006), and Poplack and Dion's (2009) real-time analysis found this effect to have already been in place by the 19th century. For the LF-origin speakers, the polarity constraint should be near-categorical and should top the ranking of variable constraints (Deshaies and Laforge [1981]; Emirkanian and Sankoff [1985]; Poplack and Dion [2009]; Poplack and Turpin [1999]; and more). In contrast, for the HF-origin cohorts, though polarity is still involved in variant selection, it is expected to have a much smaller magnitude of

effect than for LF varieties (Gudmestad et al. [2020]; Roberts [2012]; see also Villeneuve and Comeau [2016] for Vimeu French).

6.4.2. Temporal distance

Competing against the implicit polarity constraint is a factor group that represents the prescriptive rule: temporal distance. Temporal distance is consistently invoked to explain variability in the prescriptive literature according to Poplack and Dion's (2009) detailed analysis of the prescribed rules for usage of each future variant. Although inter-grammar agreement never achieves full consensus, 59% of grammars agree that the periphrastic future should be used when the event expressed by the future predication is proximate or immediate (Poplack and Dion 2009, 568). Note that neither variant is *pro*scribed in these contexts – using the periphrastic future to express distal contexts and the synthetic future to express proximal contexts is by no means 'wrong' according to normative dictate – precisely because they are thought to mean different things. Though each grammar makes it clear that these variants should *not* be treated as though they are interchangeable, the exact contexts that license the synthetic future or the periphrastic future vary from grammar to grammar. Thus, one grammar might recommend the synthetic future for distal contexts and the periphrastic future for proximal ones, whereas another might recommend the opposite. The net result of this is that the synthetic future and the periphrastic future are alternatively accepted in proximal and distal contexts – which, of course, reflects the variability that is inherent in naturalistic speech.

In contrast, variationist work has turned up only minimal effects for this proposed factor group (if any effect is found at all). Operationalizing proximity proceeds slightly differently from study to study. Poplack and Turpin (1999, 150), for example, mark all events occurring within the day as proximal and anything beyond that is coded as distal, whereas Grimm (2010, 86)

marks the cut-off from proximal to distal at one week, and Gudmestad et al. (2020) considers everything occurring within a month to be proximal. With this in mind, I initially coded for a number of fine distinctions in order to see whether they divided themselves naturally (either favouring or disfavouring one variant in particular) – within an hour (83), a day (84), a week (85), a month (86), a year (87), and longer than a year (88) – but as it turned out, temporal information was quite scarce in my data. Consequently, many of these fine distinctions were not represented in the dataset of each cohort; only the LF_1 group has the full range of distinctions. All cohorts have at least some data 'within the week' and some outside this time period, so the division between proximal and distal was established at that timepoint (which happens to correspond to Grimm's [2010] classification): events occurring within the week are coded as proximal, whereas anything occurring later than that is coded as distal.

(83) **Va être**$_{[PF]}$ là dans trente minutes. [VFC.LF_1.Bob.00/55/56]

'I'll be there in thirty minutes.'

(84) Je pars puis je **vas revenir**$_{[PF]}$ ce soir. [VFC.LF_1.Catherine.01/00/29]

'I'm leaving and I'll come back this evening.'

(85) Je **vais aller**$_{[PF]}$ la voir lundi. [VFC.LF_2.Fabienne.01/03/52]

'I'm going to see her on Monday.'

(86) L'autre **va avoir**$_{[PF]}$ uh– treize ans dans– dans deux semaines. [VFC.LF_1.Albert.00/16/08]

'My other [child] will be thirteen in two weeks.'

(87) Je viens juste de faire cinq jours à Vancouver pour un projet qui **va continuer**$_{[PF]}$ um– à l'automne. [VFC.LF$_2$.Héloïse.00/07/01]

'I just spent five days in Vancouver for a project that will continue um– in the fall.'

(88) On **dira**$_{[SF]}$ dans vingt-cinq ans si c'était juste une passe. [VFC.LF$_1$.Béatriz.00/58/59]

'We'll say in 25 years if it was just a trend.'

6.4.3. Adverbial specification

Many variationist studies have examined the effects of temporal adverbials, and one or two find that it exerts a small but statistically significant effect in LF (e.g. Poplack and Turpin 1999; Poplack and Dion 2009), but Gudmestad et al. (2020) find that it is actually the *most* important conditioning factor in HF. Differences in the effect of this factor group may (at least in part) be due to the fact that it is operationalized quite differently from study to study. Some studies simply code for the presence versus absence of an adverbial, while others incorporate information about the specificity of the adverbial into the groupings. I opted for the latter option in order to maximize comparability with studies of HF: tokens were coded as either having a specific adverbial (89), a non-specific adverbial (90), or no adverbial (91).

(89) Je **vais aller**$_{[PF]}$ la voir <u>lundi</u>. [VFC.LF$_2$.Fabienne.01/43/08]

'I'm going to see her on Monday.'

(90) On **va aller**$_{[PF]}$ au parc <u>tantôt</u>. [VFC.LF$_1$.Émmanuel.01/06/43]

'We're going to the park later.'

(91) En fait on **va aller**$_{[PF]}$ en Corse. [VFC.HF$_1$.Julianne.00/32/01]

'In fact we're going to Corsica.'

6.4.4. Style (subject type)

This factor group is included to explicitly test whether the synthetic future functions as a stylistic marker in my data. Grimm (2010) operationalizes style using topic-based coding and finds no significant effect. I instead follow other studies in operationalizing style using pronouns rather than topic, which does turn up significant effects. Poplack's work (with Turpin [1999], and later with Dion [2009]) has shown that there is a significant association between the use of the 2SG pronoun *vous* (used in polite or formal speech) and the synthetic future variant, and that this effect seems to have become stronger over time – it is more pronounced in 20th century data than it is in 19th century data. In a similar vein, Wagner and Sankoff (2011) found that more formal address pronouns *vous* and *nous* patterned together (along with nominal subjects) in favouring the synthetic future. Their analysis finds subject type to be a significant factor in synthetic future selection, and they determine that it acts as a 'proxy' for speech style in their data. Roberts (2012) finds almost an identical significant univariate effect in HF – *vous* corresponds to an increase in occurrence of the synthetic future. Thus, while this factor group does not enable us to distinguish between an HF-oriented pattern and an LF-oriented one, it is included to test the hypothesis that associations of prestige are important to selection of the synthetic future in the implicit norms of Victoria francophones.

As such, I coded for all subject types, including noun phrases (92), proper names (93), relative pronouns (94), and several types of personal pronouns (1SG *je*; 2SG *tu*; 2SG *vous* (95);

3SG *il, elle, on* (96); 1PL *nous*; 2PL *vous*; 3PL *ils, elles*), but later condensed these initial divisions according to their association with formality. the two pronouns most associated with formality (*nous* and 2SG *vous*) were coded as 'formal', all other pronouns as 'informal', and noun phrases, relative pronouns, and proper names as 'nominal'.

(92) Le problème **va être**[PF] encore de plus en plus grave. [VFC.LF$_1$.Ida.00/50/58]

'The problem will become worse and worse.'

(93) Je suis sûr que Danielle **va accepter**[PF]. [VFC.LF$_1$.Christian.01/10/57]

'I'm sure that Danielle will accept.'

(94) Je viens de finir d'écrire ce qui **sera**[SF] dans le tome deux de *Présence francophone*. [VFC.LF$_1$.Béatriz.00/26/06]

'I've just finished writing what will be in the second volume of *Présence francophone*.'

(95) « Est-ce que vous **allez parler**[PF] avec un accent québécois ou un accent français ? » [VFC.LF$_1$.Georges.00/59/44]

' "Will you speak with a Quebec accent or a French accent?"'

(96) L'anglais elle l'**apprendra**[SF] simplement avec les enfants dans– du voisinage. [VFC.HF$_1$.Fernand.00/59/33]

'She'll learn English simply with the children in the– of the neighbourhood.'

6.5. Results

6.5.1. Overall

The overall distributional results show that both the G_1 cohorts have slightly lower rates of the synthetic future variant than in studies of comparable varieties. The real difference, however, is with the G_2 cohorts, who both have very low rates of synthetic future relative to their G_1 counterparts (Table 6.1). The LF_2 cohort nearly categorically avoids the synthetic future variant, while the HF_2 cohort uses it only marginally more often – though still nowhere near the rates of either G_1 cohort or those established for HF or LF in previous studies (Table 6.2).

Table 6.1. Overall rate of synthetic future for each cohort

	%SF	N	Total
HF_1	25%	27	110
HF_2	6%	7	110
LF_1	18%	76	429
LF_2	2%	2	84
Overall	15%	112	733

Table 6.2. Overall rate of synthetic future in previous studies

LF	Author	%SF
Hawkesbury, ON	(Grimm 2010)	11%
Ottawa-Hull	(Poplack and Turpin 1999; Poplack and Dion 2009)	20%
Montreal	(Emirkanian and Sankoff 1985)	21%
Montreal	(Wagner and Sankoff 2011)	26%
HF	**Author**	**%SF**
Mainland France	(Gudmestad et al. 2020)	33%*
Mainland France	(Roberts 2012)	41%
Aix-en-Provence	(Jeanjean 1988)	58%
Paris	(François 1974)	62%

*: This study includes the present tense as a third variant. In order to improve comparability in this table, overall rate of synthetic future was calculated using the raw N of synthetic future divided by the total N of both periphrastic future and synthetic future.

Indeed, with so few tokens of synthetic future (only nine in total across both G_2 cohorts), it seems clear that the periphrastic future variant has virtually eliminated the synthetic future variant here. This means that the internal conditioning for the younger cohorts will be considerably weakened if not absent altogether, and therefore that assessing sharedness of patterns with multivariate analysis will be impossible. It is clear nonetheless that G_2 cohorts have achieved cohesion in their tendency to avoid the synthetic future. In contrast to the results for *avoir/être* where the rates of each variant were relatively uniform across all four cohorts, and in contrast to the results for assibilation where there was a clear HF/LF split, here there is a very clear G_1/G_2 split. The explanation invoked for *avoir/être* was that rates were so low across the board due to the strong negative associations with *avoir* usage, which would be exacerbated by explicit correction in

school situations for G_2 speakers, or to an existing change in progress that is already well underway in other LF-speaking communities (or possibly some combination of one enhancing the other). For this variable, the attested reduction in synthetic future aligns with the continuation of the trend away from synthetic future that characterises other LF-speaking communities[62] (cf. the real-time data in Poplack and Dion [2009]; see also Sankoff and Wagner [2020]), but examining the internal conditioning should elucidate whether associations of prestige might also play a role.

In addition, it would appear that the rates of synthetic future in both G_1 cohorts are far closer to each other than we would expect from the rates of the synthetic future in benchmark data. The rate of synthetic future in my LF_1 data (18%) is in the ballpark of expected rates for LF, given previous studies which found a range of rates from 11–26%, but the rate for HF_1 (25%) is quite low, as compared to previous studies of HF which found a range of rates from 33–62%. Although this observation in isolation does not constitute conclusive evidence of G_1 convergence, since rates can be wildly different from study to study, the assessment of underlying conditioning that follows in Sections 6.5.2–6.5.3 should confirm or disconfirm this: if conditioning differs between G_1 cohorts, then G_1 speakers can be said to have retained original usage norms, despite the apparent movement of HF_1 speakers towards the periphrastic future.

[62] Note also from Table 6.2 that the rate of synthetic future seems to be declining in HF studies as well, but since there has been no longitudinal study to confirm whether change in progress is underway in HF, I refrain from making any strong claims to that effect here. Suffice it to say that the apparent convergence of HF_1 speakers towards LF rates of synthetic future might be due to a similar change in progress away from the synthetic future that also characterises HF – or it might simply be due to the severe lack of contexts that favour its appearance (e.g. negative polarity contexts).

6.5.2. Comparison of internal conditioning for LF_1 and LF_2 cohorts

Given that LF_2 uses virtually no synthetic future at all (2%; N = 2/84), it will not be possible to compare their internal conditioning to that of LF_1, where variability appears to be quite robust. However, analysing the internal patterns for LF_1 will still be a useful exercise to determine whether they match those of noncontact benchmarks, in order to have a better idea of what input may have been provided to G_2 speakers. As we can see from Table 6.2, the rate for LF_1 resembles what has been attested for LF speakers elsewhere, and Table 6.3 confirms that their conditioning is also largely in line with expectations.

Table 6.3. Multivariate analysis of synthetic future selection for the LF_1 cohort

Corrected mean			.130
Overall %			18%
Total N			76/429
	FW	%	N
Polarity			
Negative	**.99**	92%	47/51
Affirmative	**.36**	8%	29/378
Range	***63***		
Temporal distance			
Distal	[.67]	24%	11/46
Proximal	[.33]	11%	5/46
Adverbial specification			
Specific	[.44]	17%	4/23
Non-specific	[.64]	17%	6/36
No adverb	[.49]	18%	66/370
Style (pronouns)			
Formal	[ ]	0%	0/2
Others	[ ]	18%	76/427

Not all factor groups add up to the total N because of excluded or non-collapsible categories. For instance, temporal distance could only be assessed for tokens where temporal information was available either explicitly or via context inference, which are very much in the minority in my data (N = 92).

The strength of the polarity constraint (which is virtually identical to what has been found for noncontact benchmark LF in other studies) indicates that LF_1 speakers have retained the major conditioning factor for synthetic future. Style (pronouns) is not significant here, and its effect runs in the opposite direction as expected: rather than formal pronouns favouring the synthetic future (which has been hypothesized in other studies to reflect the more positive normative evaluations associated with this variant), they never co-occur with the synthetic future in this data. However, it should be noted that there are exceedingly few tokens of formal pronouns (N = 2), which reflects the highly informal nature of the interviews. There are also indications of a slight interaction in the adverbial specification factor group, given that the percentages are mis-aligned with the weights attributed to these factors by the Goldvarb model. Cross-tabulation reveals that this factor group turns out to be non-orthogonal with the temporal distance factor group: tokens with no adverbial specification are also far more likely (categorically so, in this case) to be mute as to temporal distance. On the other hand, the factor group of temporal distance shows indications that it falls in line with noncontact LF (e.g. temporal distance), but this and the other effects just described pale in comparison to the overriding contribution of polarity, to the point of not even achieving statistical significance. In sum, these results show that the LF_1 cohort aligns with the major constraint noncontact LF, which is a good indication that systematic and implicit LF-oriented input does exist in the linguistic environment of the LF_2 speakers.

Multivariate analysis of LF_2 patterns is not possible since most speakers invariantly avoid the synthetic future, but closer examination of the individual tendencies for the two speakers that *do* make use of the synthetic future demonstrates that only one actually seems to respect the LF_1 polarity constraint, illustrated by the token in (97).

(97) Je **rentrerai**[SF] pas là-dedans là. [VFC.LF2.André.2.00/07/18]

'I won't get into it.'

However, around half the LF_2 speakers had no negative polarity contexts in their data at all, and hence they have no opportunity to demonstrate adherence to the polarity constraint, and the other half produce exceedingly few negative polarity tokens (one or two at most). In other words, the LF_2 behaviour cannot be classified as (mis)aligned with that of LF_1 – we have no way of knowing whether more data for these speakers would produce a better match with LF_1 results. With this caveat in mind, as far as we can tell for the LF_2 cohort, the periphrastic appears to be virtually the only option for expressing future time.

6.5.3. Comparison of internal conditioning for HF_1 and HF_2 cohorts

Since the HF_2 cohort also has a very low rate of synthetic future (only 6%; N = 7 tokens), it will again prove virtually impossible to glean anything reliable from multivariate analysis of their data. As such, I apply the same procedure here as I did for LF_1 and LF_2, and first assess the HF_1 patterns in order to gain a better understanding of whether HF_2 speakers are provided with systematic input. Like the LF_1 cohort, it turns out that the patterns of variant selection for HF_1 are largely in line with benchmark HF findings: polarity is nowhere near categorical, to the point of being on par with another constraint – adverbial specification – though neither are significant. The lack of significant effects is likely due to the very low token count for this cohort.

Table 6.4. Multivariate analysis of synthetic future selection for the HF_1 cohort

		HF_1	
Corrected mean			0.245
Overall %			25%
Total N			27/110
	FW	%	N
Polarity			
Negative	[.73]	50%	5/10
Affirmative	[.48]	22%	22/100
Adverbial specification			
Specific	[.47]	20%	1/5
Non-specific	[.83]	63%	5/8
No adverbial	[.47]	22%	21/97
Style (pronouns)			
Formal	[ ]	100%	3/3
Non-formal	[ ]	22%	24/107
Temporal distance			
Distal	[ ]	7%	1/14
Proximal	[ ]	0%	0/21

Not all factor groups add up to the total N because of excluded or non-collapsible categories.

By convention, ranges are not calculated for non-significant effects, but there is a clear difference between the importance of the polarity constraint for the HF_1 cohort and what was attested for the LF_1 cohort – though it should be noted that there are exceedingly few negative-

polarity contexts in the HF_1 data (N = 10), which necessarily tempers the strength of this conclusion. This suggests that the implicit norm for synthetic future-selection is different from that of the LF_1 cohort to some degree, though I hesitate to come to any strong conclusions with so few tokens – especially since no factor groups are significant. Some other interesting observations might be made with regard to this multivariate analysis, but small token counts throughout make the interpretation of these results an exercise in careful equivocation. For instance, it might be noteworthy that the effect for style (pronouns) runs opposite to that of LF_1, but with so few tokens, it is difficult to say whether this constitutes a genuine distinction between the two G_1 cohorts.

What do HF_2 speakers do, given that they presumably have a choice between the different LF_1 and HF_1 conditioning (to the extent that this was discernible from the small token counts)? HF_2 speakers avoid the synthetic future near-categorically in the same manner as we saw for LF_2, and thus have also effectively factored out any variability in this paradigm. Recall that LF_2 speakers had too few tokens of negative polarity contexts to confidently determine whether they were aligned with the LF_1 polarity constraint. The results for HF_2 speakers are similar: seven of the eight speakers in the cohort have either no or extremely few contexts of negative polarity. One speaker, Annique, has several tokens of negative polarity contexts (N = 6), but does not use the synthetic future in any of them – her single token of synthetic future occurs in an affirmative clause. Another speaker, Frédéric, has several tokens of synthetic future (N = 5), but only a single negative polarity context – in which he opts for the periphrastic future. Given the strength of the polarity constraint elsewhere, it is entirely possible that these speakers might have failed to use the synthetic future because they *coincidentally* had no negative utterances with future temporal reference. Indeed, for nearly all speakers in the HF_2 cohort, the synthetic future is

simply avoided altogether, but even speakers with higher rates of synthetic future cannot be assessed for their conformity to the HF_1 pattern, since they have very few contexts that would license its appearance. Furthermore, closer examination of the synthetic future tokens for this cohort reveals that nearly all of them appear in collocations with highly frequent and irregularly conjugated verbs (*aura*, *sera*), which suggests that these may actually function as quasi-fixed expressions (98)–(103).

(98) Ils vont apprendre mais c'est– ça– ça **sera**[SF] difficile, là. [VFC.HF_2.Frédéric.00/11/01]

'They'll learn but it– it– it'll be difficult, eh.'

(99) « Je vais rien qu'à l'école anglaise, ça **sera**[SF] facile. » [VFC.HF_2.Frédéric.00/15/59]

' "I only go to English school, it'll be easy."'

(100) On s'est dit « mieux vaut faire avant Noël, on **sera**[SF] plus confortables. » [VFC.HF_2.Frédéric.01/10/43]

'We told ourselves, "better do it before Christmas, we'll be more comfortable."'

(101) Ça **sera**[SF] peut-être plus diverse. [VFC.HF_2.Frédéric.01/13/21]

'It will maybe be more diverse.'

(102) Je me dis « donc il **aura**[SF] probablement plus de– de gens d'un peu partout. » [VFC.HF_2.Frédéric.01/13/21]

'I tell myself, "so there will probably be more people from everywhere."'

(103) Mais là à Victoria il **aura**[SF] peut-être pas le même ou une job aussi bon.
[VFC.HF$_2$.Gilbert.01/15/36]
'But then in Victoria he maybe won't have the same or a job as good.'

This makes it clear that this cohort simply does not productively make use of variable synthetic future. Rather, they show every indication of following the more general trend away from synthetic future that also characterizes the LF$_2$ cohort. In sum, what *reliable* evidence exists for both G$_2$ cohorts seems to suggest that – on the whole – they are more similar to each other than they are to either G$_1$ cohort, but also that their very low rate of synthetic future would align them more with LF benchmark data than HF benchmark data (where, as we saw, the synthetic future is maintained at a higher rate).

6.5.4. Extralinguistic factors

Fisher's two-tailed exact tests of generation show that the synthetic future is associated with G$_1$ speakers, whereas the periphrastic future is associated with G$_2$ speakers, and that this association is statistically significant (Table 6.5). There is no significant association of different origins with propensity to use one variant over another; the synthetic future forms an identical proportion of the data for LF- and HF-origin cohorts alike (Table 6.6).

Table 6.5. Cross-tabulation (2x2 contingency table) of generation and future temporal reference variants

***p* < 0.0001**	**SF**		**PF**		**Total N**
G_1	103	19%	436	81%	539
G_2	9	5%	185	95%	194
Total	**112**	**15%**	**621**	**85%**	**733**

Significance calculated using Fisher's two-tailed exact test.

Table 6.6. Cross-tabulation (2x2 contingency table) of origins and future temporal reference variants

***p* = 0.9114**	**SF**		**PF**		**Total N**
LF	78	15%	435	85%	513
HF	34	15%	186	85%	220
Total	**112**	**15%**	**621**	**85%**	**733**

Significance calculated using Fisher's two-tailed exact test.

In contrast to assibilation and *avoir/être*, where significant extralinguistic effects were sparser (they only applied to one cohort in each case), here, all cohorts but one (LF_2) are subject to a number of significant external effects. However, closer examination of the data for HF_2 revealed that each of the significant effects reflected the external characteristics of the only individual in this cohort to use any measure of synthetic future: Frédéric. As such, the focus in this section will be on the two G_1 cohorts only. For the first two variables, there was clear evidence that exposure to the broader community or maintenance of French in a range of social contexts (operationalised through a number of individual factor groups) affects the rate of variants associated with LF usage. For future temporal reference, this effect is not immediately apparent, but the results of the external factors do not contradict this general finding.

For instance, the external factor groups that affect variant selection for the HF_1 cohort are those that pertain to contexts of language usage (Table 6.7): those who use both English and French at home and those who frequently use French with friends are most likely to use the synthetic future. Since HF_1 individuals categorically have higher %HF in network (indeed, this is why this factor group is not assessed for HF_1), these results might be interpreted as evidence that HF_1 individuals who continue to engage in regular contact with other HF-origin individuals by means of their friend network are those who are most likely to retain higher rates of synthetic future. If more exposure to the predominantly LF community results in a higher rate of LF variants (as we saw for the first two variants), then it also stands to reason that if one's network is primarily HF, HF tendencies will be more likely to be retained.

Table 6.7. Bivariate analysis of extralinguistic effects on synthetic future selection for the HF_1 cohort

	SF		**Tot N**
	25%	27	110
Language in home	***p* = 0.002**		
Mostly French	12%	7	58
Both	39%	20	52
Language with peers	***p* = 0.003**		
Mostly French	33%	24	72
Both	8%	3	38

Light gray denotes significant favouring effects.

Table 6.8. Bivariate analysis of extralinguistic effects on synthetic future selection for the LF$_1$ cohort

	SF		Tot N
	18%	76	429
Exposure to school Fr.	***p* = 0.008**		
Higher	28%	24	85
Lower	15%	52	344
Restriction	***p* = 0.052**		
None	20%	65	324
Moderate	9%	7	76
Severe	14%	4	29
Infrastructure contact	***p* = 0.077**		
Higher	21%	42	197
Lower	15%	34	232
%Life in Victoria	***p* = 0.008**		
High	8%	7	93
Moderate	22%	37	172
Low	20%	32	164

Light gray denotes significant favouring effects.

The external effects for the LF$_1$ cohort are somewhat different (Table 6.8). Individuals with a higher degree of contact with community infrastructure, those who have lived in Victoria for a smaller proportion of their lives, and those who are unrestricted tend to favour the synthetic future. On the one hand, this suggests that maintenance of the synthetic future as a variant within this system depends on regular usage of French (or relatively less time in Victoria – though recall that even the latest arrivals in my sample have lived in Victoria for at least a decade). On the

other hand, the broader LF community is increasingly moving towards the periphrastic future, but there is no guarantee that older speakers are able to participate in this change, given the relative inflexibility of their linguistic systems compared to younger speakers. In other words, use of the synthetic variant by G_1 speakers might simply reflect the fact that they are less able to modify their speech patterns even after sustained periods of contact. An alternative explanation for the persistence of the synthetic future at G_1 is the association of prestige with this form – this is supported by the significant effect of exposure to school French on the synthetic variant in the LF_1 cohort. Even though neither variant in the future temporal reference sector is explicitly stigmatised, there are clearly positive implicit associations with the synthetic future within this community, which might prevent the synthetic future from losing too much ground to the periphrastic future. If this is the case, however, then why do the G_2 speakers not also persist in using the synthetic future? In addition to hearing the synthetic variant spoken by G_1 speakers, they receive explicit instruction in conjugation and usage of the synthetic variant from grade 3 onward (*Ministère de l'Éducation de la Colombie-Britannique* 2010, 174). It is not the case that G_2 speakers receive so little exposure to this variant that they might be considered unaware of it. The more likely explanation is that they are aligned with the change in progress towards the periphrastic variant, but that the general lack of synthetic-favouring contexts for both G_2 cohorts has possibly caused the G_2 rate of the synthetic variant to appear especially low. At any rate, this does not change the finding that pertains to the main goal of this study: G_2 speakers behave cohesively, despite clear differences in pattern (and, to a lesser extent, rate) at G_1.

6.6. Summary and interim conclusions

There are at least two major trends that characterize the future temporal reference variable in Victoria. Firstly, the synthetic future is governed by different internal patterns for G_1 speakers of

different origins, though those for HF$_1$ did not emerge as significant. LF-origin speakers' use of the synthetic future is primarily governed by the major polarity constraint, which is the only internal factor to emerge as statistically significant. In contrast, no factors are significant in the HF$_1$ data, but polarity plays a less categorical role for HF$_1$ and is at least on par with – if not outranked by – adverbial specification, which is in line with noncontact HF. The second major finding of this chapter is that G$_2$ speakers *do* share rates for future temporal reference, which happen to be in line with the overarching change away from synthetic future that characterises LF elsewhere. The synthetic future is avoided quite consistently by G$_2$ speakers across the board regardless of origins – with the exception of one HF$_2$ speaker, who tends to use the synthetic future in contextually restricted expressions like '*ça sera*'.

Recall from the preceding chapter that, although *avoir/être* showed a similar reduction in rate across the board, it was not clear whether this reduction was somehow linked to a stigma against *avoir*, or whether it was simply the continuation of a change in progress away from *avoir* that characterises other LF communities – though of course, it could also be both in tandem – since both the stigma and the change in progress would result in less *avoir*. The new evidence from future temporal reference allows this to be retroactively disambiguated to a certain degree, since the G$_2$ speakers do not use elevated rates of the synthetic future (which would be consistent with an explanation rooted in positive associations with this variant). Indeed, they barely use the synthetic variant at all! This can only be explained by the change in progress underway in other LF communities, with the extreme avoidance of the synthetic variant likely due to a general lack of synthetic-favouring contexts for both G$_2$ cohorts.

Chapter 7. Yes/no questions: A three-way conflict with salient variability

The conclusions drawn thus far from the analysis of the first three variables are that G_1 speakers largely tend to maintain their distinct original norms, whereas G_2 speakers tend to gravitate towards the LF norm both in terms of rates and (where this is possible to assess) underlying patterns. This was demonstrated uncontroversially for assibilation and *avoir/être*, though *avoir/être* was also subject to an additional effect from a change in progress underway in benchmark LF, likely coupled with (or exacerbated by) the explicitly-taught prescriptive injunction against the use of *avoir*, which acted to depress the overall rate of *avoir* for all cohorts. The effects of future temporal reference were somewhat more indeterminate, but did not contradict this general conclusion. G_1 speakers seem to have rather similar rates of the synthetic future, though a difference is maintained between HF_1 and LF_1 as far as the ranking of internal constraints that favour this variant. G_2 speakers were found to align with expectations from benchmark studies as far as the documented change in progress towards the periphrastic future. There was evidence to suggest that speakers might retain the synthetic future in quasi-fixed expressions with irregular verbs, but there were too few negative polarity contexts to confidently verify their adherence to this major LF constraint. Taken together, the results from these three variables suggest that G_2 speakers converge on LF-oriented norms (to the extent that this is observable – no variables are *in*consistent with this), but that G_1 speakers may only deviate from benchmark patterns for their respective varieties in terms of rates, and only if stigma against the form in question motivates them to do so (cf. *avoir* avoidance). I now undertake the final variable of this study, yes/no questions, which instantiates a conflict between the LF usage model, the HF usage model, and grammatical prescription simultaneously, but which also involves a

greater number of variants and conditioning constraints as compared to any of assibilation, *avoir/être*, or future temporal reference.

7.1. Background

To form yes/no questions (sometimes termed *polar* or *total* questions), French offers four different options[63]: rising intonation with subject-verb word order (INT; (104)), the grammaticalized interrogative expression *est-ce que* (ECQ; (105)), subject-verb inversion (INV; (106)), and the post-verbal interrogative particle *-tu*[64] (TU; (107)).

(104) Donc toi **t'as fréquenté**[INT] l'ancienne École-Brodeur? [VFC.LF$_1$.Albert.00/39/33]

'So you went to the old École-Brodeur?'

(105) **Est-ce que c'est**[ECQ] le climat qui attire les gens? [VFC.LF$_1$.Albert.00/50/17]

'Is it the climate that attracts people?'

(106) **Vas-tu**[INV] encore être ici en 2018? [VFC.LF$_1$.Albert.02/19/54]

'Are you still going to be here in 2018?'

[63] A fifth variant, complex inversion, has been effectively out of usage since the 18th century and is therefore not considered here (Elsig and Poplack 2006).

[64] Though most often pronounced [ti] in HF – apparently on the model of grammatically 'correct' constructions ending in *-t-il(s)* (e.g. *c'est-il*, pronounced [seti]; [Grevisse and Goosse 2016, 539]), in LF it is categorically pronounced [ty]. Note also that this is *not* inversion: the following particle is invariantly *-tu* regardless of grammatical person of the subject.

(107) Tu te **souviens-tu**$_{[TU]}$ de cette histoire-là? [VFC.LF$_1$.Albert.01/37/25]

'Do you remember that drama [lit. 'story']?'

These four options are selected at different rates and according to different conditions in the implicit usage norms of HF and LF (discussed in more detail below). In addition, the prescriptive and pedagogical recommendations for this variable diverge quite considerably from both of these usage norms. Specifically, they endorse variants that are quite rare in natural speech, such as inversion or *est-ce que* (Grevisse and Goosse 2016), arguing for their widespread acceptability in any context despite the rarity of their actual usage or their restriction to a narrow functional domain. No doubt as a result of these prescriptions, many pedagogical sources present inversion and *est-ce que* as though they are the only options for question formation (Antes 2016; Elsig 2009; Etienne and Sax 2009). For example, inversion is confined to highly formal contexts in naturalistic spoken HF (elaborated in more detail below), but this construction is privileged across the board in French textbooks – "often inappropriately, from a sociolinguistic standpoint" (Etienne and Sax 2009, 597), since it is frequently used in examples where a casual or more intimate variant (per the implicit usage norm) would be more suitable. The example reproduced below evokes quite a formal tone via its selection of inversion (among other elements – note the use of the formal address pronoun *nous*, which further adds to the impression of elevated tone), which is at odds with the conversational scenario being depicted (a child addressing her parents):

(108) [Reproduced from Etienne and Sax 2009, 591, Excerpt 2]

Alice, qui n'a que quatre ans, est très curieuse. Elle pose des questions sur tout ! Pour aider ses pauvres parents, trouvez une réponse logique à chacune de ses questions.

Alice : Pourquoi **voulons-nous**[INV] éviter les heures de pointe?

'Alice, who is only four, is very curious. She asks questions about everything! To help her poor parents, find a logical answer to each of her questions.

Alice: Why do we wish to avoid rush hour?'

Conversely, normative sources disparage or ignore the use of other variants that are very common in actual usage – if they are even mentioned at all. For example, though the *-tu* particle is a variant frequently used in naturalistic spoken LF, it almost never features in prescriptive or pedagogical literature, or is mentioned merely in passing as a variant associated with LF. It is not generally recommended as an option for question formation in school French. Such omissions are emblematic of the normative enterprise, which may create significant confusion for the student of French grammatical rules. In other words, yes/no questions are a variable where a significant portion of the implicit usage norm is either ignored or outright condemned by the prescriptive sources – unlike future temporal reference, where both variants are perfectly acceptable in the eyes of grammarians and teachers (though, as we saw in the previous chapter, they are used according to the implicit usage norm, *not* the prescriptive rules of school French).

7.2. Benchmarks

7.2.1. Normative French

Generally speaking, prescriptive manuals tend to lag behind usage in both the acknowledgement and – especially – the *endorsement* of variant forms (cf. Poplack et al. 2015). In the case of yes/no questions, this has meant that inversion has long held the title of preferred (and in many cases only[65]) variant mandated in the prescriptive literature, regardless of its actual frequency in HF or LF speech (cf. Elsig 2009, 29). The most recent edition of *Le bon usage* (Grevisse and Goosse 2016) manages to concede that inversion seems to be confined to "careful language, mostly in writing" (*ibid*., 538); in contrast, intonation is acknowledged as the most frequent option in speech[66], albeit one that is explicitly classed as informal or popular (*ibid*., 548). Note that intonation is not deemed to be structurally 'deficient' or 'illogical' in any way (a strategy frequently employed by traditional grammars to discount a particular variant; see Poplack et al. [2015, 36]); it is simply attributed to a less formal register (another common prescriptive strategy for 'explaining' variation).

Similarly, the latest edition of *Le bon usage* also acknowledges the use of the *-tu* variant in LF. Where historically *-tu* has been classed as 'vulgar' at best (cf. Elsig 2009, 29[67]), it is now described somewhat more neutrally by Grevisse and Goosse (2016, 539) as "still thriving in

[65] As Elsig (2009, 28) puts it, "[g]rammars from before the first quarter of the twentieth century treat pronominal inversion as if it were the only way to form a question."

[66] Note that manuals such as *Le bon usage* are intended as reference books for any and every variety of French, but that the authors' point of reference for specific comments on usage is undoubtedly HF. As such, their statement on intonation can be understood as 'INT is the most frequent option *in HF*.'

[67] Also 'popular', 'rural', and 'ridiculous' (Elsig, 2009, 29).

certain regions". In other words, although they qualify it as decidedly evoking popular speech and certainly do not explicitly endorse its widespread use outside these regions, any overt *stigma* that is associated with its use in France may be waived for speakers of LF in Canada. This more neutral recent assessment of *-tu* usage by *Le bon usage* is not echoed in grammar manuals or pedagogical texts, however: Etienne and Sax's (2009) review of lessons on interrogative phrases in 22 French textbooks turned up no mention of *-tu* at all. The textbooks they examined were college-level French textbooks intended to instruct L2 speakers in French usage; as such, though representative of the sorts of constructions included in pedagogical literature for 'advanced' learners of French, they likely do not *directly* correspond to the materials in use by French teachers in the L1 classroom. However, as it turns out, omissions of *-tu* are common in pedagogical literature intended for use in majority-context, L1 classrooms as well. Of particular note are documents such as *Enseigner la grammaire actuelle: Programme de recherche en littératie*, a 2019 Government of Quebec publication (Centre de transfert pour la réussite éducative du Québec 2019) intended to provide pedagogical guidance to teachers in 'modern' grammar, but which only includes examples of *est-ce que* and inversion in their module on interrogative structures – despite *-tu* being an important and even emblematic variant of the LF interrogative system:

(109) [Examples reproduced from CTREQ [2019, 9]

Est-ce que$_{[ECQ]}$ Margot court derrière ses frères?

Margot **court-elle**$_{[INV]}$ derrière ses frères?

Does Margot run behind her brothers?

In sum, inversion has been the prototypically acceptable form suggested by traditional grammars over the last several centuries (cf. Elsig and Poplack 2006, 83), and contemporary pedagogical sources continue to privilege use of inversion and *est-ce que* as question-formation strategies while completely ignoring or outright condemning the other two variants, intonation and – to a far greater extent – *-tu*. This is the case even in material that purports to capture actual usage (e.g. CTREQ 2019). In some grammars, the multiple variants of yes/no questions are brought into line by attributing each variant to a different communicative function. According to the latest edition of *Le bon usage* (Grevisse and Goosse 2016) the selection of each variant would depend on the register of a given speech context: inversion continues to be heavily promoted as a generally acceptable strategy (it is frequently used to illustrate 'general' examples of interrogations, cf. Grevisse and Goosse [2016, 528, 532]), but tends to be associated with formal speech contexts; intonation is privileged in speech, a generally informal context; *-tu* is only acceptable in certain geographic regions (including Canada) and is suggested to be receding as a direct result of schooling (in a similar way as was suggested for *avoir*).

(110) On inversion:

> (528) Par la phrase interrogative, on demande une information à l'interlocuteur :
> *Gèle-t-il ? À quelle heure dinez-vous ?*
>
> 'By means of the interrogative phrase, one requests information from their interlocutor: Is it freezing out? At what time do you dine?'

(532) L'interrogation globale appelle une réponse par *oui* ou par *non* (ou *si*). Elle peut porter sur le verbe : *Pleut-il ?* – ou sur le verbe en relation avec un autre élément : *Pleut-il souvent ? Votre frère accepte-t-il ma proposition ?*

'Total interrogation demands a response with yes or with no. It can be applied to the verb: Is it raining/Does it rain? – or to the verb in relation to another element: Does it rain often? Will your brother accept my proposition?'

(538) La langue soignée, surtout écrite, marque l'interrogation par l'inversion du sujet.

'Careful language, particularly written, marks interrogation with the inversion of the subject.'

(111) On intonation:

(548) Ce procédé est devenu la forme ordinaire de l'interrogation dans l'oral quotidien.

'This procedure has become the ordinary form of interrogation in everyday speech.'

(112) On *-tu*:

(539) Quoiqu'encore bien vivant dans certaines régions (comme en Normandie et au Québec), le tour paraît aujourd'hui en recul, sans doute sous l'influence de l'école.

'While still alive and well in certain regions (such as Normandy and Quebec), this turn of phrase appears to be receding these days, no doubt under the influence of the school.'

I have intentionally left the *est-ce que* variant out of the above summary, because – in contrast to the straightforward inclusion of *est-ce que* in pedagogical texts – the prescriptive recommendations for *est-ce que* (both historical and modern) are considerably more complex than those accorded to the other variants. As of the most recent edition of *Le bon usage*, *est-ce que* is recognised as being far more frequent than inversion – even on par with intonation in popular speech:

> Dans la langue parlée courante, cette construction est beaucoup plus fréquente que l'inversion…mais elle est très fortement concurrencée par l'interrogation marquée par l'intonation seule. (Grevisse and Goosse 2016, 544)
>
> 'In popular speech, this construction is far more frequent than inversion, but it is very strongly contested by interrogation marked by intonation alone.'

However, it has historically been subject to vacillations in acceptability, with the *Académie française* (the preeminent authority on the French language) "unreservedly" recommending the use of *est-ce que* until 1935 (Grevisse and Goosse 2016, 545), but then reversing their recommendations in 1987 to *pro*scribe its use:

> En 1987, *art. ce*, rompant avec ce qu'elle avait écrit antérieurement, l'Acad. critique les ex. mêmes qu'elle donnait comme normaux en 1932: « [...] *On doit dire Quand partirez-vous ? et non Quand est-ce que vous partirez ? À qui dois-*

> *je m'adresser ? et non À qui est-ce que je dois m'adresser ?* (Grevisse and Goosse 2016, 545).
>
> 'In 1987, art. ce, breaking with what it had written previously, the *Académie* criticises the same examples that it gave as normal in 1932: "We must use When will you leave$_{[INV]}$? and not When will you leave$_{[ECQ]}$? To whom must I address$_{[INV]}$ myself? and not To whom must I address$_{[ECQ]}$ myself?'

Grevisse and Goosse's (2016, 545) contemporary evaluation of *est-ce que* appears rather contradictory as well, in that they maintain that it is considered "inelegant", but is also well-attested in even the "noblest of genres", as well as in "elaborate literary language" over the last few centuries (e.g. in the celebrated works of Vaugelas, Maupassant, Pascal, and Renard). The association of *est-ce que* with these great authors would normally legitimize it in the eyes of the prescriptive enterprise, but it is nonetheless evaluated as awkward or 'heavy' by several grammarians (e.g. Chevalier et al. 1973; Dauzat 1943; Gaiffe et al. 1936; Gougenheim 1962), 'colloquial' (Foulet 1921) or somehow brusque (Terry 1970) by others, and even 'substandard' in earlier editions of *Le bon usage* (Grevisse 1993). Despite these conflicting evaluations, it is still the *only* question-formation variant other than inversion to be explicitly *recommended* by prescriptive literature – albeit in the specific context of first-person subjects only, where inversion is quite infelicitous (except in certain lexical verbs, such as *puis-je*, cf. Elsig [2009, 19]) and where *est-ce que* is therefore deemed to be the acceptable substitute for inversion (Chevalier et al. 1978; Gaiffe et al. 1936; Grevisse and Goosse 2016, 538, 546).

If *est-ce que* is considered appropriate (in certain contexts), why is it evaluated so disfavourably by so many grammarians? Possibly at the root of distaste for *est-ce que* is its

"redundancy" or "incongruence" with predicative clauses, as suggested by several authors (see Druetta [2003, 27]; Elsig [2009, 31]). For example, *est-ce que c'est*, prior to the historical grammaticalization of *est-ce que*, would have literally translated to 'is it that it is', yielding a 'redundant' double predicate. However, given that *est-ce que* has not been considered a productive conjugation of the verb *être* for many centuries, the redundancy argument for the avoidance of *est-ce que* in modern texts seems altogether quite weak. Another possible source of prescriptive aversion to *est-ce que* is that its form is subject to considerable variability in WH constructions, resulting in variations of *est-ce que* such as '*où ce que c'est qu'il va ?*' or '*où qu'il va ?*'[68] that are strongly stigmatised ("*franchement vulgaires*," Valdman [2000, 660]). In contrast, *est-ce que* categorically retains its full form /ɛsk(ə)/ in yes-no questions (Tailleur 2013, 135)[69]. Grevisse and Goosse – somewhat more gently than Valdman – recommend that 'double' *est-ce que* be confined to the familiar register only.

> Le double introducteur *est-ce que c'est que* appartient seulement à la langue parlée familière. (Grevisse and Goosse 2016, 547)
>
> 'The double introducer est-ce que c'est que [is it that it is] belongs only to familiar spoken speech.'

[68] Examples from Valdman (2000, 658).

[69] In fact, Tailleur (2013, 135) argues that '*est-ce que*' is at different stages of grammaticalization in yes/no questions and in WH-questions – with WH-*est-ce que* being much further along than YN-*est-ce que* (see also Elsig [2009, 192]).

In other words, the apparent vacillation in ECQ's prescriptive acceptability may actually stem from the different evaluations accorded to reduced or reduplicative forms of ECQ, which happen to only surface in WH questions, and to full-form ECQ, which is categorical in yes/no questions. If this is the case, it seems likely that *est-ce que* specifically in *yes/no questions* (the focus of this chapter) can be ranked alongside intonation and inversion as an acceptable, non-stigmatised option. This classification of (yes/no) *est-ce que* as fundamentally acceptable is supported by its frequent inclusion in pedagogical texts, as outlined above.

The use of each variant according to stylistic context is something that is empirically testable using spontaneous speech data. Less testable, however, are claims such as Dauzat's (1943; see also Mosegaard Hansen, [2001, 471]), that *est-ce que* conveys a more 'emphatic'[70] question than intonation or Mosegaard Hansen's that *est-ce que* implies "a stronger degree of doubt about the proposition than V-Cl [INV]" (2001, 464). The intention of such claims is presumably to differentiate the use of *est-ce que* from INT, which are the two yes/no question variants that are possible in informal contexts according to the prescriptive rule – again, in an attempt to attribute forms to specific functions in a one-to-one relationship. However, although prescriptive literature

[70] If 'emphasis' can be interpreted to mean 'prosodic prominence', then arguably it would be possible to use certain acoustic cues to distinguish emphatic from non-emphatic utterances. Indeed, there are many studies that have examined the relationship between prosody and question formation (e.g. Beyssade 2006; Cheng and Rooryck 2000; Déprez, Syrett, and Kawahara 2012; Hamlaoui 2011), but these studies do not suggest that prosodic prominence is a contributing factor to the selection of one question-formation strategy over another within the same variable context (though there may be differences in, for instance, the laboratory-elicited prosody of WH and yes/no questions; cf. Déprez, Syrett, and Kawahara [2012]). Furthermore, given the vagaries of phonetic production in spontaneous speech (cf. Poplack et al. 2020), it is unlikely that acoustic measures of focus would even be reliable, since various parts of a phrase might receive prosodic prominence for reasons completely unrelated to emphasis.

often refers to a speaker's internal motivation or emotional state to distinguish the use of one variant or another, such determinations rely on the subjective interpretation of the author and are therefore generally too unreliable for our purposes. As such, the register dimension is tested in this study as a possible factor influencing the selection of yes/no question variants, but the 'emphasis' dimension is not.

7.2.2. Implicit usage norm in LF

All four variants of yes/no questions (cf. (104)–(107)) are present to some degree in the LF implicit usage norm, but their distribution is not perfectly even. While *-tu*, inversion, and intonation seem to be used at roughly equal rates (each around a third of the time, per Elsig and Poplack [2006], Elsig [2009], and Fox's [1991] studies), *est-ce que* is selected far less frequently overall – as little as 0.6% of the time in Fox's (1991, 281) Saint-Sauveur, Quebec data – and is favoured by stylistic factors that suggest a strong influence of speech register. For instance, Elsig (2009, 89, 95) finds *est-ce que* to be strongly favoured by formal speech styles and conspicuously absent from other linguistic contexts – to the point that he does not include *est-ce que* in his discussion of "productive" variants in LF (*ibid*., 165). In other words, the *est-ce que* variant clearly has its function in the implicit LF norm – as a "prestige marker" of formal style (*ibid*., 105) – but this function (and, concomitantly, the *est-ce que* form) seems to surface very rarely in speech.

Aside from the link between *est-ce que* and formal speech contexts, previous variationist studies have uncovered additional trends in internal conditioning for the three other yes/no question variants in the implicit LF usage norm (Elsig 2009; Elsig and Poplack 2006; Fox 1991). For instance, Elsig (2009, 175) finds that inversion – though used at a frequency roughly equal to that of the other variants in his data – is favoured by informal speech contexts, most often co-

occurs with five highly frequent verbs (*savoir*, *vouloir*, *voir*, *avoir*, and *être*), and is limited to second-person subjects. He concludes that inversion in LF has *functionally* receded to this narrow linguistic context of occurrence despite its relatively frequent use (*ibid*., 130, 189; Fox 1991, 282). He also argues that the *-tu* variant shows signs of having encroached onto inversion's former functional territory[71]. Elsig's (2009, 94) evidence for this is that, in LF, *-tu* is favoured with rare or uncommon verbs, as well as with subject types other than second-person. The occurrence of *-tu* in a wider range of linguistic contexts, as well as its specific association with LF varieties of French, leads him to suggest that it has become the 'default' yes/no question-formation strategy in LF (*ibid*., 90, 105, 145). The remaining variant, intonation, is by far the predominant variant selected in negative polarity contexts[72] (Elsig and Poplack 2006, 4), but is otherwise the "least constrained by independent variables" overall (Elsig 2009, 91). It is very slightly favoured in more formal speech styles (*ibid*., 89), which contradicts the prescriptive suggestion that it be confined to informal contexts.

In all, there are five main diagnostics gleaned from the conditioning of yes/no question variants in LF that can be used to determine whether G_2 speakers are aligned with the LF implicit usage benchmark or not: (a) the use of inversion, *-tu*, and intonation at roughly equal rates, but *est-ce que* at a very low rate, (b) the restriction of inversion to 2P subjects, possibly in addition to its predominance in certain frequently-occurring lexical verbs (*savoir*, *vouloir*, *voir*, *avoir*, and

[71] According to Elsig's (2009) syntactic analysis, the structural similarities between inversion and *-tu* may have allowed the latter to supplant the former. I refer to reader to his detailed analysis for more information, since it is beyond the scope of this dissertation to address this claim here.

[72] This invariance in negative polarity contexts leads the authors to exclude this context from further study. Results from both Elsig (2009) and Elsig and Poplack (2006) are on affirmative polarity tokens only. This is elaborated further in Section 7.3.

être), (c) the restriction of *est-ce que* to formal speech styles, (d) the use of *-tu* in a wide range of linguistic contexts but preferred in informal speech, and (e) the (subtle) association of intonation with more formal speech. In addition to these main diagnostics, there are a number of more ancillary trends uncovered by previous variationist analysis of LF data that may be useful in clearing up whether G_1 speakers are, in fact, providing 'full' vernacular input to their G_2 counterparts. However, as these are far subtler than the main diagnostics described above, they are less useful in clearly distinguishing LF from HF or from prescriptive rules and shall therefore only be adduced if needed in the coming sections.

7.2.3. Implicit usage norm in HF

The HF usage norm contrasts markedly with that of LF in both the rates of each variant and the patterns that govern their use. Firstly, whereas LF has a slight imbalance in the distribution of each variant within the yes/no question system – in that it uses all four variants, but not equally often (*est-ce que*), nor equally productively (inversion) – the variant distribution in HF is even more imbalanced. The implicit norm for HF largely prefers intonation: as Elsig (2009, p. 15) puts it "intonation questions are clearly the default variant in the context of yes/no interrogatives" in spoken HF, and this finding is replicated across a number of studies (e.g. Coveney 2002; Pohl 1965). As for the other variants, *est-ce que* is far rarer in speech (cf. Coveney 2002, 190), inversion is considered to be effectively unproductive in speech, instead being confined to written materials (e.g. letters [Pohl 1965]; stage-play scripts [Terry 1970]) except for certain formulaic utterances (e.g. '*comment vas-tu?*'), and the *-tu* variant is not used at all.

Although the interrogative system in HF has been extensively studied over several decades, there is considerable disagreement in the literature about which factors affect which variants. The effect of subject type is one such area of disagreement: Terry (1970, 90) finds that inversion

seems to frequently occur with second-person subjects, but Behnstedt (1973, 152) finds that these contexts actually seem to *dis*favour inversion in his data. This illustrates a key difference between the implicit norm of LF and that of HF, since inversion is found to categorically occur with 2P subjects in LF[73]. Furthermore, Elsig (2009, 130) notes that while inversion in LF is restricted to 2P subjects, inversion in HF has been *lexically* restricted to a handful of effectively fixed constructions since at least the 15th century – though there is disagreement as to which constructions those might be (e.g. *as-tu*, *avez-vous*, *est-il*, *vois-tu*, *voyez-vous*, *penses-tu*, *comprenez-vous*, *voulez-vous*, *veut-il*, *pouvez-vous*, *peut-on*, etc.; examples drawn from Ashby [1977)]; Behnstedt [1973]; Coveney [2002]; Pohl [1965]), which might at least partly be due to the fact that some of the comparison studies predate Elsig's by several decades.

The expectation for LF is that inversion will *categorically* occur with 2P subjects, but as the literature for HF shows, there is no such expectation for HF, and examples such as *est-il* or *peut-on* clearly demonstrate INV's admissibility in lexical and pronominal contexts outside the narrower one attributed to LF. Furthermore, recall that the rate of inversion is very different in LF and HF, with inversion used at roughly the same rate as intonation and *-tu* in LF, but clearly a very minor variant in HF speech (Pohl 1965). In other words, inversion is frequent but tightly constrained to a narrow pronominal context in LF, whereas it is far rarer but used in a wider number of pronominal contexts in HF. Finally, inversion is favoured by *in*formal contexts in LF, but in HF seems to be largely associated with *formal* contexts (Behnstedt 1973; Coveney 2002;

[73] Elsig (2009, 77) lists a single exception to this tendency – one token with 3P PL *ils* – a token which he then goes on to demonstrate as highly unusual in other respects. His conclusion is that "inversion in yes/no questions [in LF] is categorically excluded from all non-second person contexts."

Pohl 1965), suggesting that it has received opposite sociostylistic evaluations in the implicit norms of HF and of LF.

The behaviour of *est-ce que* in HF is also quite distinct from its behaviour in LF – at least on the surface – since it is attested as the second-most frequent yes/no question variant in HF (though *well* behind intonation nonetheless), but is by far the least-attested variant in LF. There is also no evidence in the published literature on yes/no questions to suggest that HF privileges the use of *est-ce que* over inversion with first-person subjects – as recommended by prescriptive rules for certain lexical verbs. For example, Terry (1970, 90) finds that 1P pronouns occur at equal rates with both *est-ce que* and inversion – but his study is not designed to determine whether 1P pronouns particularly *favour* or *disfavour* these constructions. Behnstedt (1973, 152), on the other hand, finds that 1P subject pronouns have neither a favouring nor a disfavouring influence on the choice of inversion. There is also little evidence to suggest that *est-ce que* in HF speech is confined to formal contexts, though many authors seem to evaluate it as somewhat *less* colloquial than intonation (Coveney 2002; Terry 1970), with other studies qualifying this variant as downright familiar – some even going so far as to class it as 'uneducated'[74] (see Elsig 2009, 27). *Est-ce que* in HF is clearly not confined to the same hyperformal context as in LF.

The preceding sections have shown that HF and LF differ materially in their variable selection of yes/no question variants, and that both varieties diverge from the prescriptive rules. Furthermore, several differences between HF and LF (as well as between LF or HF usage and

[74] It is difficult to tell whether assessments of this type refer to *est-ce que* in yes/no questions specifically, or to *est-ce que* in WH-questions as well, which tend to receive more stigma because of their relative variability (cf. '*où ce que c'est qu'il va ?*' and the other "*franchement vulgaire*" examples from Valdman [2000, 660]), since many authors conflate the two constructions in their analyses.

school French) seem to be along stylistic lines, with the interrogative variants occurring at different points on the informal-formal scale in each of the two varieties. Other diagnostics, such as the restriction of inversion to 2P subjects in LF, are structural in nature. The diagnostics that will be applied in the analysis are summarized in Table 7.1.

Table 7.1. Summary of diagnostics used for yes/no questions

	Implicit norm: HF	**Implicit norm: LF**	**School French**
INT	Highly frequent Informal contexts	Approx. 1/3 of tokens More formal contexts	Informal contexts
ECQ	Less frequent Neutral	Rare Hyperformal contexts	Use with 1P subjects
INV	Rare Formal contexts 'Fixed' expressions	Approx. 1/3 of tokens Informal contexts 2P subjects only	Acceptable everywhere
-TU	N/A	Approx. 1/3 of tokens Informal contexts	N/A

7.3. Variable context and exclusions

Scanning through the list of variants in (104)–(107), it is clear that these are all ways of asking a yes/no question, but it is less clear for some authors whether all these questions are *equivalent* in a linguistic sense. In his detailed analysis of the communicative functions of interrogatives, Coveney (1996, 123–75) argues that requests for information, opinion, clarification, advice, action, or permission are all different in terms of communicative function (determined according to over two-dozen features pertaining to the relationship between the speaker and addressee; the

knowledge, assumptions, or beliefs of the question's author; propositional content of the question; etc.), and claims that these functions apparently take different interrogative structures – some categorically so. Other authors proceed in a similar manner (e.g. Mosegaard Hansen 2001).

It is of course possible to examine one's finite dataset for functions, and to retroactively conclude that whichever (categorical or variable) forms populate a given function are the *only* forms that this function allows, as Coveney has done (1996, 178–80). However, post-hoc reconstructions of this sort are notoriously problematic. First of all, it is often the case that the taxonomy of functions itself is not identical across datasets. Several different functions in one study may be subsumed under a single function in another (Coveney 1996, 150); various functions may be sorted into completely different categories from study to study (Coveney 1996, 151); or if one function is mentioned in two different studies, it may not be the case that the same interrogative structures are used for this particular function. Even within a single study, attempts to achieve inter-analyst reliability in determining a particular function are not always successful – Coveney himself notes that the native-speaker judges used in his study occasionally had "[an] interpretation of the communicative function of [an] utterance that differed from [his] own" (Coveney 1996, 183).

Another issue with this technique is that it flies dangerously close to the prescriptive doctrine of 'form-function symmetry' whereby variation is entirely factored out by attributing each variant form to a very specific function[75]. That is, as the functions get more and more specific, the forms that populate these functions become more 'categorical' – so long as the analyst makes sure to include dozens of qualifications, amendments, or exceptions to these supposedly

[75] The pitfalls of this are also clearly illustrated in Poplack and Dion (2009) for the future temporal reference sector, as we saw in the preceding chapter.

categorical rules[76]. Furthermore, if one particular interrogative structure is found to be categorical in a dozen different functions (as is the case for SV interrogatives with intonation in Coveney [1996]) it obscures the fact that this particular structure is not truly determined by any given communicative function, but rather is more of a default interrogative structure in *any* function (cf. Elsig and Poplack 2006, 82).

Like in other variationist work on yes/no questions (e.g. Elsig 2009; Elsig and Poplack 2006), in this study I abstract away from referential equivalence that depends on nuanced semantic or subjective reconstructions and instead confine my investigation to 'questions where a positive/negative answer is expected' – regardless of what might be the underlying intent of the speaker or the interpretation by the listener. Thus, rhetorical questions (113) are excluded, because a positive/negative answer is not *expected* (though one might nonetheless be given). Other types of questions are excluded because the full spectrum of variability is not available for some reason: tag questions (114); questions asked in English (115); questions without an inflected verb (116); indirect interrogatives (117); and direct echoes of the interviewer (118).

(113) **Tu sais je veux dire**? [VFC.LF$_1$.Ida.00/56/49]

'You see what I'm saying?'

(114) C'est en février, **hein**? [VFC.LF$_1$.Ida.00/42/16]

'It's in February, huh?'

[76] For example, in Coveney (1996, 178–79): where "requests for action which take the form *tu/vous* + non-modal verb in the present tense are categorically SV...[but] if a modal verb is present or if the verb is in the future tense, then all three variants of YNQ [yes/no questions] are possible."

(115) But do they get colds? [VFC.LF$_2$.Jenna.01/58/13]

(116) **Moi**? Ah non, non, moi j'ai jamais connu ça. [VFC.LF$_1$.Donald.00/04/06]

'Me? Ah no, no, I've never known that.'

(117) Je sais pas **si ça vous arrive aussi**? [VFC.HF$_1$.Julianne.00/09/26]

'I don't know if that happens to you also?'

(118) [Suzanne] Non c'est le contenant. [Catherine] **C'est le contenant**?

[VFC.LF$_1$.Catherine.02/23/58]

[Suzanne] 'No it's the container.' [Catherine] 'It's the container?'

Since each of the contexts included in a variationist study must license all the possible variants, it was also important to exclude negative polarity contexts at the outset. Previous studies have found that some yes/no question variants are effectively prohibited in negative polarity contexts (e.g. *-tu*, *est-ce que*; cf. Morin [2017]; Tailleur [2013]), whereas others are strongly or even categorically preferred (intonation; cf. Elsig [2009]; Elsig and Poplack [2006]). In HF, Mosegaard Hansen (2001, 198) finds near-categoricity of intonation in negative polarity contexts. The categoricity of intonation in negative polarity contexts also occurs in LF; in fact, since polarity is such a major determinant of variant selection, the variationist analyses in Elsig and Poplack (2006) were carried out on affirmative polarity tokens only.

Some authors have suggested that this imbalance in the variants licensed in negative polarity contexts is due to the syntactic particularities of interrogative structures and negation in French (which I will not elaborate in any great detail, though see Morin [2017] for a comprehensive analysis). Söll (1983) notes that utterances with negative polarity using *est-ce que* are judged less

'harmonious' than those using INT[77]. Morin (2017, 176) (citing Vinet [2001]) claims that "negative [*-tu*] questions in French are only possible if the preverbal negative marker *ne* is present." Similarly, Coveney (1996, 212) asserts that "it is generally true today that *ne* is obligatory in [inversion] interrogatives." If these statements are all true, then per Söll (1983) and Mosegaard Hansen (2001), we would expect no *est-ce que* in negative-polarity utterances, and per Morin (2017) and Coveney (1996), we would expect no inversion or *-tu* where the negation marker *ne* is absent. Given that 99.8% of negative utterances in French *lack* this negation marker (Poplack and St-Amand 2007, 723), this would leave intonation as essentially the only option for utterances with negative polarity. Note, however, that Morin, Coveney, Söll, and Mosegaard Hansen to a greater or lesser degree all rely on grammaticality judgments to come to their conclusions about the incompatibility of negation and certain interrogative structures. While this is not always an *in*appropriate strategy, using it to come to firm conclusions about variant selection is risky since such analyst or native-speaker judgments are often very subjective. As a case in point, where empirical evidence exists, it corroborates some of these intuitive, judgment-based impressions but it contradicts others. For example, Elsig and Poplack's (2006) study of yes/no questions in LF found that utterances with negative polarity are "overwhelmingly expressed by intonation" (*ibid*., 86). It might appear, then, that suggestions that inversion and *est-ce que* are avoided in utterances with negative polarity because they are less 'harmonious' are true (cf. Söll 1983). On the other hand, Elsig and Poplack (2006) also find that a smaller but significant proportion of negative polarity utterances are expressed by *-tu* in their data, directly

[77] "*Notre informateur français [...] acceptait toujours les substitutions du type* tu connais pas? → est-ce que tu connais? *mais refusait* *est-ce que tu connais pas?" (Söll 1983, 48). See also Mosegaard Hansen (2001, 428).

contradicting Morin's (2017) claim that *-tu* is only possible when the negation marker *ne* is present. If we were to rely solely on grammaticality judgments to conclude that some variants are not permitted in negative polarity contexts, we might erroneously rule out the occurrence of *-tu* in such contexts. It is therefore important to first objectively verify within the usage data whether negative polarity *does* co-occur with more variants than intonation – even if, for some scholars and their informants, these constructions may be judged as 'inharmonious'.

In this study, polarity was coded as negative if there were any elements indicating negation (including *ne*, *n'*, *pas*, *jamais*, etc.) in the interrogative clause (119), or affirmative if the clause lacked such negation elements (120).

(119) **Vous avez** pas **vu**$_{[INT]}$ ce film? [VFC.HF$_1$.Julianne.00/28/11]

'You haven't seen that film?'

(120) **Est-ce que vous avez vu**$_{[ECQ]}$ le– le film 'Bienvenue chez les Ch'tis'?

[VFC.HF$_1$.Julianne.00/28/07]

'Have you seen the film Bienvenue chez les Ch'tis?'

To ascertain whether the polarity effect is present in my data, I calculated the proportion of negative and affirmative tokens used in each of the four variants for each of the four cohorts to determine whether negative polarity utterances categorically occurred with the intonation variant. As expected, I found that the negative polarity effect holds for all four cohorts (Table 7.2): negative contexts are predominantly (even *categorically*, for the HF$_1$ corpus) populated by tokens of the intonation variant – though note that the effect seems slightly weaker in the G$_2$ cohorts

(and weaker still in LF_2 than in HF_2, mirroring the differences between the G_1 counterparts for each of these cohorts). All cohorts (except HF_1) also have very restricted variability in the negative polarity contexts, with only *est-ce que* as a rare alternative to intonation[78]. Since the polarity effect seen here is never explicitly mentioned in grammars or pedagogical manuals, its presence in all cohorts here is initially suggestive of an adherence to usage norms, rather than to prescriptive rules.

Table 7.2. Effect of affirmative (+) and negative (–) polarity on yes/no question variant selection

	HF_1		LF_1		HF_2		LF_2	
	+	–	+	–	–	–	+	–
	84%	16%	92%	8%	92%	8%	97%	3%
TU	—	—	22%	0%	—	—	21%	0%
ECQ	32%	0%	32%	5%	50%	11%	38%	25%
INV	3%	0%	10%	0%	5%	0%	4%	0%
INT	65%	**100%**	36%	**95%**	45%	**89%**	38%	**75%**

We can observe from this table that polarity is not diagnostic of a difference between HF and LF, since in both varieties negative polarity seems to strongly promote intonation, but it clearly distinguishes usage norms from the rules of school French – nowhere in the pages of *Le bon usage* does it stipulate that intonation must be used in negative polarity utterances. Note that HF_1

[78] Note that its use in negative polarity contexts here is extremely minimal, with all cohorts except HF_1 using *est-ce que* only once. However, this does represent a contrast with Elsig and Poplack's (2006) study, where *-tu* was the only attested alternative to intonation in negative polarity utterances.

speakers have almost double the proportion of negative polarity contexts in their data compared to the other three cohorts (compare HF_1's 16% negative polarity contexts to 8% in LF_1 and HF_2, and 3% in LF_2). Thus, removing negative polarity contexts may have a greater effect on the variant distribution in the HF_1 cohort specifically, by more dramatically depressing the rate of intonation (thereby also inflating the rate of other variants) in HF_1 than in the other cohorts. This is an important consideration for when we review the patterns within the HF_1 cohort in coming sections.

In sum, the near-categoricity of intonation in negative polarity contexts precludes them from variable analysis. Affirmative polarity tokens, on the other hand, admit a wider range of variation (as was also the case in Elsig and Poplack's [2006] and Mosegaard Hansen's [2001] studies) and as such, the remaining analyses will focus on tokens in these contexts only. Once all of the exclusions are accounted for (some 350 tokens), my dataset contains 522 tokens distributed over the four different yes/no question variants shown in (104)–(107).

7.4. Hypotheses about variant selection

As mentioned in the sections on implicit usage norms in HF and LF above, what makes the interrogative system such an appealing variable for my research question is the fact that both its variant distribution and its conditioning are demonstrably different in HF and in LF (Elsig 2009; Elsig and Poplack 2006). Furthermore, each of these usage norms has areas where they apparently diverge from the rules of school French: for instance, both usage norms use intonation almost exclusively for negative polarity tokens, neither usage norm seems to privilege the use of inversion (though LF uses it with some frequency, this usage is confined to a narrow context of occurrence and therefore is not fully productive), nor do they confine their use of *est-ce que* to 1P subjects. In addition, neither of the usage norms aligns one-to-one with the sociostylistic

continuum that has been proposed by the prescriptive enterprise: LF does not confine its use of inversion to formal contexts or its use of intonation to informal ones, and also makes liberal use of *-tu*; HF uses intonation in informal contexts, and splits formal contexts between intonation and *est-ce que*.

These differences will make it possible to compare the underlying patterns in the different cohorts of speakers and to draw conclusions about whether they have come to share implicit usage norms, as well as whether these norms align clearly with one of the three possible sources of influence. This variable has been of great interest to many scholars, with the result that several internal factors have already been proposed (though not all have been tested in a variationist framework), but the present study is confined to examination of those factors which are clearly and robustly diagnostic of differences between the HF and LF implicit usage norms.

7.4.1. Subject type

Subject type has been found to be a key factor group in the occurrence of the inversion variant: in LF, inversion categorically occurs with 2P subjects, while in HF no such restrictions apply. Although the crucial factor seems to be 2P versus all other subject types, I follow Elsig (2009, 54–55) in initially coding for each different type of subject that occurred in my data, including 1P SG (121), 2P SG (122), 3P SG (*il*, *elle*) (123), *on* (refers to *nous*) (124), 2P PL (125), 3P PL (126), expletive (*il*, *c'*) (127), nominal NP (128), and non-clitic demonstrative (*ça*, *cela*) (129) subjects.

(121) J'**ai-tu**$_{[TU]}$ vraiment besoin de lire tout ça? [VFC.LF$_1$.Albert.00/05/36]

'Do I really need to read all this?'

(122) Puis lui il m'avait dit : « tu m'**as-tu vu**[TU]? » [VFC.LF1.Catherine.02/04/44]

'And he told me "did you see me?"'

(123) Il **était-tu**[TU] meilleur, lui? [VFC.LF1.Donald.01/22/44]

'Was he better?'

(124) **Est-ce qu'**on[ECQ] **peut** aller prendre le souper avec Camille?

[VFC.LF2.Fabienne.01/04/26]

'Can we go have dinner with Camille?'

(125) **Est-ce que**[ECQ] vous **êtes** pour Trump? [VFC.HF1.Hélène.01/17/48]

'Are you for Trump?'

(126) **Est-ce qu'**ils[ECQ] **ont** une ferme? [VFC.HF2.Annique.00/16/43]

'Do they have a farm?'

(127) **Est-ce qu'**il[ECQ] y en **a** d'autres? [VFC.HF2.Daniel.00/05/25]

'Are there any more?'

(128) Ça **va-tu**[TU] marcher? [VFC.LF1.Catherine.01/33/02]

'Will it work?'

(129) **Est-ce que**[ECQ] ton chum est italien alors? [VFC.HF1.Hélène.01/10/08]

'Is your boyfriend Italian then?'

7.4.2. Factor groups related to lexical identity of verb

Some authors specifically mention lexical identity of the verb as having an effect on variant selection, while others operationalize factor groups that are in some way related to lexical identity, such as verb frequency (cf. Elsig 2009). The hypothesis that underlies factor groups like these is that certain interrogative structures may behave as more or less fixed lexical chunks – particularly in those variants that are thought to be receding (e.g. inversion, which seems to only be maintained in a handful of quasi-fixed lexical expressions, such as *avez-vous*). In some cases, lexical identity and frequency can be very difficult to disentangle, in the sense that a 'highly frequent' category may be populated by only one or two lexemes – in which case it becomes difficult to determine whether the lexeme itself or its high frequency is what contributes to variant selection (though see Walker [2012] for a skilful manipulation of this issue). Recall that previous work in the yes/no question paradigm has shown that, in HF, the inversion variant seems to occur more frequently with a limited set of rather frequent verbs (Dewaele [1999, 174] lists *avoir*, *pouvoir*, *trouver*, *faire*, and *vouloir*) or only in a few lexically-fixed expressions (Behnstedt [1973, 21, 52] mentions *vois-tu*/*voyez-vous*, *veux-tu*/*voulez-vous*, and *as-tu*/*avez-vous*). However, in LF, Elsig's (2009, 78–79) detailed analysis of the effects of lexical verb on variant selection shows that many more verbs than the ones listed above seem to favour inversion, including many verbs that he identifies as being in some way related to 'cognitive' states or activities (*avoir, trouver, voir* and *vouloir* are among these, but there are many others as well). Elsig concludes that while inversion is not necessarily as productive as, say, *-tu* in Laurentian French, it is not truly lexicalized or fixed in the way that it seems to be for HF either. In order to test whether inversion is used in frequent verbs, I calculated a corpus-internal measure of frequency, with each verb lexeme being coded as 'rare' (10 or fewer occurrences), 'very

infrequent' (11-20 occurrences), 'infrequent' (21-50 occurrences), 'frequent' (51-100 occurrences), 'very frequent' (101-500 occurrences), and 'extremely frequent' (501-7,000 occurrences). The initial frequency divisions were somewhat too severe for the amount of available data, and so were later condensed into 'high' (>501 occurrences) and 'low' (<500 occurrences).

7.4.3. Style (topic)

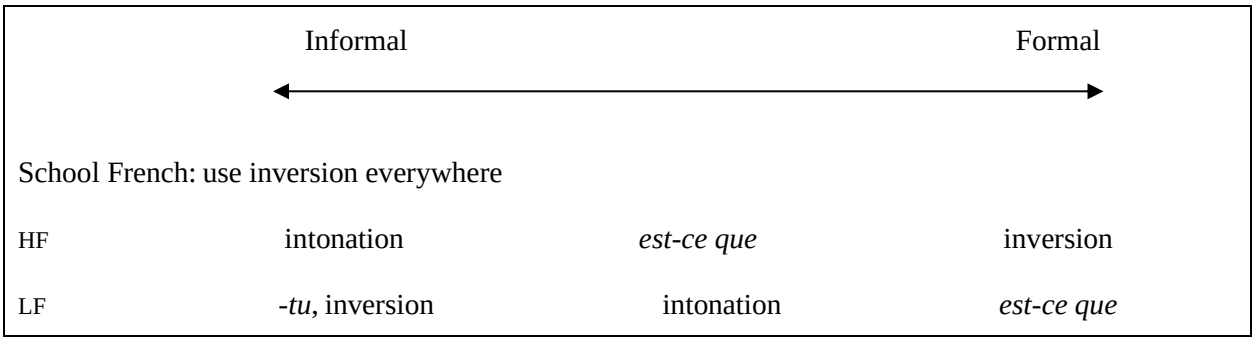
A very common strategy in prescriptive literature is to attribute different variants to different speech styles or registers (see Poplack et al. [2015] for detailed examples). The interrogative system is no exception, with linguists and grammarians alike proposing that the continued use of several variants in this paradigm is because each variant occurs along a sociostylistic continuum (Chevalier et al. 1978; Dauzat 1943; Elsig 2009; Foulet 1921; Gaiffe et al. 1936; Gougenheim 1962; Grevisse and Goosse 2016; Valdman 2000). The exact position of each variant on the continuum is not identical from author to author, although there is widespread agreement on the position of certain variants, such as *-tu*, which is almost always placed at the informal extreme of the scale (if it is mentioned at all), and *est-ce que*, which tends to be situated towards the formal end.

There are a number of areas where normative French, HF usage, and LF usage diverge in this respect. Although traditional grammar tends to evaluate inversion as being the formal option (cf. Section 7.2.1), this is not reflected in the implicit norm of LF, where inversion is favoured in *in*formal contexts. Conversely, where intonation tends to be classed as informal in both prescriptive literature and the HF implicit norm, intonation seems to be slightly favoured in *formal* contexts in the LF implicit usage norm. The sociostylistic classification of *est-ce que* is a bit of a quagmire, as we saw in Section 7.2.1, which I tentatively attributed to the contrast

between 'acceptable' yes/no *est-ce que*, which is categorically produced as /ɛsk(ə)/, and 'unacceptable' variations of WH-*est-ce que* (cf. Valdman's [2000] "*franchement vulgaire*" forms, such as '*où ce que c'est qu'il va ?*'). Since many studies lump together yes/no and WH- versions of *est-ce que*, both these types might therefore tend to receive a blanket evaluation as 'sub-standard' (cf. Mosegaard Hansen 2001). As long as these two *est-ce que* types are distinguished, however, it is generally agreed that yes/no *est-ce que* is fairly neutral in style (cf. Antes 2016; Etienne and Sax 2009; Fox 1991, 280). Elsig and Poplack (2006, 82) found its use to be strongly correlated with formal speech style (for LF): formal speech styles favour the use of *est-ce que* while informal speech styles disfavour it (see also Fox [1991, 282]). Bigot's (2008, 232) results on *Radio-Canada* French strongly corroborate this classification: he reports 79% *est-ce que* in the context of radio interviews. Villeneuve (2017) also finds a high rate of *est-ce que* in a different corpus of formal Quebec French.

In sum, the findings for style in the LF implicit usage norm suggest that *-tu* and inversion should be situated at the informal end of the scale, with *est-ce que* at the opposite end, and intonation erring on the formal side of neutral. In contrast, HF situates intonation at the informal end, inversion at the formal end, and *est-ce que* somewhere in the middle. The prescriptive norm prefers inversion across all contexts. Figure 7.1 summarizes the ranking of variants along a sociostylistic continuum for each of the varieties under comparison here.

Figure 7.1. Stylistic expectations for yes/no questions in school French and in the implicit usage norms of HF and LF



7.5. Results

7.5.1. Overall

In Section 7.2 above, I outlined several diagnostics within the yes/no question system that can be used to determine whether G_2 speakers conform to LF- or HF-oriented usage norms, or whether they instead converge on the rules provided by school French. These diagnostics hinge on the overall preponderance of each variant relative to the others, on their occurrence in casual or careful speech style, as well as on their co-occurrence with 2P and 1P pronouns (for inversion and *est-ce que*, respectively).

The overall rate of each variant (Table 7.3) is a first indication that HF_1 and LF_1 speakers each conform to the implicit usage norm associated with their respective benchmark varieties (see Table 7.4), and that G_2 speakers largely share the same tendencies – at least as far as the (diagnostic) proportional distribution of each variant within the whole yes/no question system.

Table 7.3. Overall distribution of yes/no question variants for each cohort

	TU		ECQ		INV		INT		TOTAL
	N	%	N	%	N	%	N	%	N
HF_1	—	—	23	32%	2	3%	47	**65%**	72
HF_2	—	—	55	**50%**	5	5%	49	**45%**	109
LF_1	51	**22%**	73	32%	23	**10%**	83	36%	230
LF_2	23	**21%**	42	38%	4	4%	42	38%	111
Overall	74	14%	193	37%	34	7%	221	42%	522

Table 7.4. Overall distribution of yes/no question variants in previous studies

LF	TU	ECQ	INV	INT
(Elsig and Poplack 2006)	33%	6%	26%	35%
(Fox 1991)[a]	34%	1%	29%	36%
HF	TU	ECQ	INV	INT
(Mosegaard Hansen 2001)	—	4%	4%	92%
(Coveney 1996)	—	21%	—	79%
(Söll 1983)	—	8%	1%	91%
(Ashby 1977)	—	11%	9%	80%
(Terry 1970)	—	3%	11%	86%
(Pohl 1965)	—	14%	1%	86%

[a]: Percentages calculated from the average of Fox's Saint-Sauveur and Sainte-Foy data.

The HF_1 cohort is well-aligned with benchmark HF usage norm data in terms of rates: they predominantly use intonation, sometimes use *est-ce que*, very rarely use inversion, and never

use *-tu*. The LF_1 cohort uses a wider range of variants distributed more evenly – as is expected from LF benchmark usage data – though their use of both *-tu* and inversion is somewhat lower than benchmark rates. As for the G_2 speakers, we see that the proportions of each variant are very similar in both cohorts. Both make little use of inversion, but use *est-ce que* and intonation at roughly equal rates – they only differ in their use of *-tu*.

This suggests two lines of division within this community: speakers differ according to origin *and* according to generation. The distributional results show which cohorts favour which variants, and an HF/LF split is obvious. Both the LF-origin cohorts favour the use of *-tu*, whereas both the HF-origin cohorts favour the use of intonation – each of those variants corresponding to an informal (arguably even emblematic) variant in the respective benchmark usage norms of these varieties. In each of the other variables, it has been clear that G_1 speakers maintain dialect heterogeneity, but G_2 speakers – while clearly aligning with other G_2 speakers rather than with G_1 speakers – were difficult to categorize as LF- or HF-oriented for the other variables (except assibilation), because their use of variants was too categorical to assess their underlying patterns. Here, we see the first clear indications that HF_2 and LF_2 have each retained some of the distinct patterns associated with the G_1 norms, though it also seems clear from these results that the G_2 cohorts share variant distributions to a large extent.

Although there are elements that divide G_2 speakers according to the patterns of their respective G_1 counterparts, they also cannot be said to be completely distinct from each other. For one thing, both G_2 cohorts share roughly the same variant distribution (with the exception of their different treatment of *-tu*), but this appears to have come at the expense of clearly aligning with either of the LF or HF benchmark proportions of each variant. By this I mean that HF_2 has decreased their use of intonation relative to HF_1, which also makes them slightly divergent from

benchmark HF usage data, and LF_2 has decreased their use of both inversion and *-tu* relative to LF_1 and to benchmark LF usage data. In addition, both G_2 cohorts have huge increases in their use of *est-ce que* relative to noncontact benchmark data. In other words, each of the G_2 cohorts has decreased the variant that – in their usage benchmark – is associated with informal speech (intonation for HF_2 and inversion/*-tu* for LF_2), and increased their use of a variant that is normatively acceptable in both (*est-ce que*). What's more, it looks like this process has parallels at G_1: while the HF_1 cohort's use of *est-ce que* might be considered roughly in line with that of benchmark HF (in the sense that it comes second to intonation, but is more frequent than inversion), the LF_1's cohort is most certainly *not*. We turn now to comparisons of underlying conditioning for each of the cohorts to continue to assess the extent to which each diagnostic reveals something about G_2 alignment.

7.5.2. Comparison of internal conditioning for LF_1 and LF_2 cohorts

Given the relatively small total token count for this variable (N = 522) and the comparatively large number of variants, it is highly likely that there will be empty or very small cells in the analysis of each diagnostic factor group. This might affect whether certain factors are selected as significant or not. However, recall from the discussion in Section 3.2 (cf. Poplack and Tagliamonte 2001, 92) that the relative *ordering* of factor groups (the constraint hierarchy) is more important than whether each of these is selected as significant – especially in cases where there are poorly distributed cells or low token counts.

The multivariate analysis for the LF_1 cohort confirms that it largely aligns with benchmark LF patterns (Table 7.5; significant effects are in bold font).

Table 7.5. Internal conditioning of yes/no questions for the LF$_1$ cohort

	TU			ECQ			INV			INT			
Corrected mean			0.222			0.287			0.045			0.362	
Overall %			22%			32%			10%			36%	
Total N			51/230			73/230			23/230			83/230	
	FW	%	N	FW	%	N	FW	%	N	FW	%	N	Tot N
Subject type													
2P	[.46]	19%	24	**.42**	25%	31	**.82**	18%	22	[.52]	38%	48	125
Others	[.55]	25%	26	**.60**	41%	42	**.14**	1%	1	[.48]	33%	34	103
Range				***18***			***68***						
Verb frequency													
Highest	[.51]	23%	31	[.50]	33%	44	[.63]	13%	17	**.45**	31%	42	134
Low	[.49]	21%	20	[.50]	30%	29	[.33]	6%	6	**.58**	43%	41	96
Range										***13***			
Style													
Casual	[.42]	17%	16	**.40**	21%	19	**.52**	11%	10*	**.64**	51%	47	92
Careful	[.57]	28%	32	**.58**	36%	41	**.49**	10%	11	**.38**	27%	31	115
Range				***18***			***3***			***26***			

*: Not all factor groups add up to the total N because of excluded tokens or non-collapsible categories.

The diagnostics we must examine in this table include the categorical association of inversion with 2P pronouns and certain lexical verbs, as well as the association of each variant with either careful or casual style. While inversion occurs with quite a wide variety of lexical verbs here (11

different lexemes, both frequent and infrequent, though *vouloir*, *pouvoir*, *être*, and *avoir* combined make up well over half the data), Table 7.5 shows that inversion is very strongly associated with 2P pronouns – only one token of inversion occurs with a non-2P pronoun.

(130) Non, parle-moi de– un petit peu de la– l'autre association, y **a-t-il**$_{[INV]}$ un conflit? Ou est-ce que vous travaillez ensemble? [VFC.LF$_1$.Christian.00/26/09]

'No, tell me about– a little bit about the– the other organization, is there a conflict? Or do you work together?'

However, as example (130) shows, this particular token is stilted and very formal in tone, and therefore does not really reflect 'typical' usage at all. In benchmark LF usage data (e.g. Elsig 2009; Elsig and Poplack 2006), inversion is not used for formal utterances, *est-ce que* is. In other words, this odd token of inversion does not detract from the finding that inversion effectively categorically occurs with 2P pronouns.

There is also a significant effect of style for all variants except *-tu*: *est-ce que* is favoured in careful speech, whereas both inversion and intonation are favoured in casual speech. Style is the major conditioning factor for *est-ce que* but a minor one for inversion, whose use is primarily governed by the 2P constraint. This aligns well with benchmark usage data. In contrast, intonation has been shown in previous studies to be associated with careful, rather than casual, style in LF, and this is not borne out by the results here. The behaviour of *-tu* with regard to style is also somewhat odd. It is not a statistically significant finding, but the multivariate analysis suggests that *-tu* is favoured in *careful* speech – which is entirely unexpected given the strong association of this variant with informal, vernacular LF. Looking over the LF$_1$ tokens of *-tu* that

occurred in careful style reveals that this variant is frequently used when speakers are engaged in complaining – for example, about certain politicians (131), inattentive students (132), or language attitudes (133).

(131) Glen Clark, **ça** te **rappelle-tu**$_{[TU]}$ des mauvais souvenirs ça? [VFC.LF$_1$.Albert.01/37/12]

'Glen Clark, does that bring back bad memories?'

(132) **Ils ont-tu**$_{[TU]}$ **compris** la moitié? Le quart? Le tiers? **Ils ont-tu**$_{[TU]}$ **compris** quelque chose? [VFC.LF$_1$.Catherine.01/19/47]

'Did they understand half? A quarter? A third? Did they understand anything?'

(133) C'est quoi parler bien le français? **Il** y en **a-tu**$_{[TU]}$ une façon de parler bien le français? [VFC.LF$_1$.Béatriz.01/04/54]

'What is speaking good French? Is there one way to speak good French?'

While these are certainly 'formal' topics per the coding protocol established in Section 5.4 (see Table 5.1 on page 145), they do not seem to instantiate a particularly formal style of speaking. If anything, speakers who are airing their grievances might not be self-monitoring at all – if unmonitored speech is likely to occur during emotional narratives of personal experience (cf. Labov 2001), why shouldn't it also occur during emotional expressions of annoyance? In contrast, the favouring effect of careful style on *est-ce que* in LF$_1$ *does* seem to reflect an elevated or more formal tone, as demonstrated by the examples in (134)–(136).

(134) **Est-ce que**[ECQ] ça– ça va entraver le déroulement d'une permanence francophone? [VFC.LF$_1$.Béatriz.01/00/42]

'Will it hinder the development of francophone maintenance?'

(135) **Est-ce que**[ECQ] ça te ferait plaisir? [VFC.LF$_1$.Christian.00/18/24]

'Would it make you happy?'

(136) **Est-ce que**[ECQ] vous accepteriez que ce soit en chinois? [VFC.LF$_1$.Catherine.01/39/18]

'Would you accept that it be in Chinese?'

Since *est-ce que* is far more frequent in the LF$_1$ data than expected from benchmark distributions, and since it clearly retains some association with formal speech style in LF$_1$, it might be the case that the high proportion of *est-ce que* in LF$_1$ speech is actually due to an overrepresentation of formal tokens. Table 7.5 shows that careful tokens are more frequent than casual ones in LF$_1$ data, and (as we will soon see from the results for the other cohorts) they are the only cohort for whom this is the case. As such, the high rate of *est-ce que* in LF$_1$ speech might at least partly be due to their elevated proportion of careful speech for this variable.

The shift in conditioning for the intonation variant (from being favoured in formal styles in benchmark LF, to being favoured in informal styles in my data) seems to align the LF$_1$ cohort with the conditioning associated with HF. Intonation has also picked up conditioning based on lexical frequency of the verb: the least frequent verbs favour intonation, whereas the more frequent verbs disfavour it. In Elsig's (2009) work, favouring effects of casual styles and of lower-frequency verbs are interpreted as markers of 'default' status – except that in Elsig's study, this qualification was associated with the *-tu* variant, not with intonation. To borrow

Elsig's interpretation here would be to suggest that LF$_1$ speakers are supplanting *-tu* with intonation as the default variant. There is good reason to believe that they might be motivated to do so, since *-tu* may have some negative associations – or at least strongly connotes 'vernacular' LF – and we have seen from other variables (e.g. *avoir/être*) that there is a tendency to restrict use of variants with such associations. Furthermore, the attitudes expressed by my participants about LF (cf. (6)–(8) on page 24) suggest that features which are emblematic of vernacular LF might end up carrying some stigma in the Victoria context.

As such, instead of being the 'default' informal variant, *-tu* in LF$_1$ might instead be reserved for more emotional or demonstrative speech (e.g. in expressions of annoyance), where a more 'colourful' variant might be deemed more acceptable. In contrast, the very high rate of *est-ce que* might be due to the 'formal' space left behind by the conversion of intonation to 'default' variant: since intonation is no longer associated with formal speech, *est-ce que* is left with more functional work to do, and consequently surfaces more often in speech. Either way, it seems clear that LF$_1$ speakers are (either deliberately or subconsciously) employing higher rates of variants that have strong positive associations.

In a similar fashion, the proportional distribution of the different variants shown in Table 7.3 suggests that LF$_2$ might not align with LF benchmark data either – for example as far as the reduction in inversion and *-tu*, and the dramatic increase in *est-ce que* – but the underlying conditioning shows that they are actually more aligned with benchmark LF than they initially appear (Table 7.6; significant effects are in bold font). For instance, LF$_2$ speakers have adopted the categorical constraint that inversion must be used with 2P pronouns – even despite the much lower rate of inversion compared to both LF$_1$ and to benchmark LF usage data – and they also use inversion with lexical verbs that are strongly inversion-selecting in benchmark data, such as

vouloir or *pouvoir* (*es-tu* [N=1], *veux-tu* [N=1], *peux-tu* [N=2]), though the data here is rather too sparse to draw any strong conclusions.

Table 7.6. Internal conditioning of yes/no questions for the LF_2 cohort

	TU			ECQ			INV			INT			
Corrected mean			0.201			0.375			0.036			0.371	
Overall %			21%			38%			4%			38%	
Total N			10/111			42/111			4/111			38/111	
	FW	%	N	FW	%	N	FW	%	N	FW	%	N	Tot N
Subject type													
2P	[.38]	14%	5	**.64**	51%	18	[ ]	11%	4	**.34**	23%	8	35
Others	[.56]	24%	18	**.44**	32%	24	K.O.	0%	0	**.58**	45%	34	76
Range				***20***						***24***			
Verb frequency													
High	[.51]	22%	18	[.47]	33%	27	[.43]	2%	2	[.54]	42%	34	81
Low	[.47]	17%	5	[.59]	50%	15	[.68]	7%	2	[.40]	27%	8	30
Style													
Casual	[.60]	26%	18	[.50]	40%	27	[.55]	4%	3	[.44]	29%	20	68
Careful	[.32]	11%	4	[.51]	39%	14	[.41]	3%	1	[.62]	47%	17	36

They also seem to have adopted some novel conditions: both *-tu* and intonation are favoured by 'other' subjects (though non-significantly so for *-tu*), whereas *est-ce que* is favoured by 2P subjects. Furthermore, although none of the style conditioning came out as significant for LF_2, the

direction of effect seen here reveals that *-tu* and inversion are associated with casual contexts, whereas intonation is associated with careful contexts. This reflects LF benchmark usage conditioning quite closely, though – remarkably – it differs from LF_1 style conditioning, which as we saw appears to have undergone some reanalysis. This is a noteworthy finding: where did LF_2 pick up benchmark LF conditioning if not from LF_1? It is possible that the external conditioning analyses will provide more conclusive answers, but I propose some initial conjectures here.

The LF_1 speakers clearly provide sufficient input to the LF_2 speakers for all four variants of this system, since they have adopted all four variants. There is no way for LF_2 speakers to acquire *-tu* from a grammar book; their use of this variant can only come from exposure to usage. However, equally clear is the fact that LF_2 speakers have struggled to acquire the *functional* conditioning for variant selection for all but the *most* categorical constraint (confining the use of inversion to 2P subject contexts). In contrast, the *sociostylistic* conditioning can be thought of as reflecting the conditions in which LF_1 speakers use these variants – which happens to align with benchmark LF even though LF_1 linguistic behaviour does not. For instance, LF_1 speakers save *-tu* for more emotional speech, which the topic-based coding protocol classes (somewhat inaccurately, in this case) as 'formal'. Emotional speech could easily be interpreted by LF_2 speakers as a 'casual' context (i.e., one in which it is acceptable to use 'vernacular' LF variants), and therefore contribute to their association of *-tu* with a more casual register.

Alternatively, their retention of LF-oriented style conditioning for intonation (favoured in careful style) suggests that intonation has not taken on a 'default' quality, whereas I proposed that it had done so in LF_1. This would appear to be because *est-ce que* has stepped into the 'default' role for this cohort: it is favoured by less frequent verbs (cf. Elsig's interpretation) and is completely neutral with respect to style at G_2. In LF_1, the retention of the style constraint on

ECQ might block it from being adopted as the 'default' variant, so they opt for the next most neutral option (i.e., intonation), but the complete unimportance of style to *est-ce que* selection at LF_2 would allow it to take on the 'default' role. In other words, LF_2 has clearly acquired many elements of the underlying grammar of LF, even where LF_1 may have abandoned them – arguably for reasons rooted in relative stigma or prestige of the variants involved.

7.5.3. Comparison of internal conditioning for HF_1 and HF_2 cohorts

Unlike what we saw for LF_1, the HF_1 conditioning does not directly contravene what we would expect from benchmark HF data (Table 7.7). The rates closely resemble benchmark expectations: the *-tu* variant is not used at all, and the HF_1 system is effectively divided between intonation and *est-ce que*. Constraints on intonation include a favouring effect of 2P subjects, of higher-frequency verbs, and of casual style. Of these, only the effect of casual style corresponds directly to benchmark expectations, since functional conditions on intonation were not well-attested in the literature. As for *est-ce que*, it is favoured by 'other' subjects, lower-frequency verbs, and careful style. Of these, the style effect and – possibly – the effect of 'other' subjects (cf. the somewhat nebulous effect of 1P subjects on *est-ce que* attested from the literature on HF) both reflect benchmark conditions. In other words, HF_1 speakers have remained relatively well-aligned with the HF benchmark.

Table 7.7. Internal conditioning of yes/no questions for the HF_1 cohort

	TU			ECQ			INV			INT			
Corrected mean			x			0.283			x			0.683	
Overall %			0%			32%			3%			65%	
Total N			0			23/72			2/72			47/72	
	FW	%	N	FW	%	N	FW	%	N	FW	%	N	Tot N
Subject type													
2P				**.33**	18%	7	[ ]	5%	2	**.65**	76%	29	38
Others				**.69**	47%	16	K.O.	0%	0	**.34**	53%	18	34
Range				***36***						***31***			
Verb frequency													
High				[.37]	27%	13	[ ]	2%	1	**.59**	71%	35	49
Low				[.51]	43%	10	[ ]	4%	1	**.31**	52%	12	23
Range										***28***			
Style													
Casual				**.37**	21%	8	[ ]	3%	1	**.62**	76%	29	38
Careful				**.67**	43%	13	K.O.	0%	0	**.35**	57%	17	30
Range				***30***						***27***			

The effect of style shows that intonation is reserved for more casual speech – which was adduced as a diagnostic between LF and HF patterns, though the reader will recall that this was one area where the LF_1 patterns did not align with the LF benchmark. The interpretation applied for this constraint in LF_1 was not 'convergence with HF_1' – which would be quite difficult given

the barriers to interaction between LF_1 and HF_1 speakers – but rather that intonation is being reanalysed as the 'default' variant for LF_1 (instead of more stigmatised *-tu*), thereby coincidentally aligning it with HF_1.

As for HF_2, the underlying system looks to have parallels with both LF_1 and HF_1 – though admittedly this is tenuous since none of the factors came out as significant for this cohort (Table 7.8). There is the suggestion of a division of labour between intonation and *est-ce que* in a similar manner as we saw for HF_1, since 2P subjects favour INT, while others favour *est-ce que*. However, this cohort uses slightly more inversion than HF_1, and like LF_1, their use of inversion is confined to 2P subjects and the verbs *être* and *pouvoir*. There are some similarities in the conditioning between HF_2 and LF_2, as well – particularly as concerns the complete lack of a style effect for *est-ce que*, which distinguishes the G_2 cohorts entirely from either of the G_1 cohorts. In the LF_2 cohort, the likelihood of *est-ce que* occurring in careful and casual tokens was close to 50/50, and we see the same trend here for HF_2.

In light of the observations made above, this seems to confirm that *est-ce que* has stepped into the G_2 grammar as the default question-formation option. Though the variants that have apparently been reduced in each of the G_2 cohorts are not 'marked' in the sense of 'stigmatised' (save perhaps for *-tu*), the variant that is being increased (*est-ce que*) has unquestionable normative appeal. In other words, instead of unidirectionally adopting LF rates and patterns, which has been the case thus far, we see more of a 'common target' approach whereby both LF_2 and HF_2 abandon some of the conditioning provided by the G_1 cohorts, but gain new (shared) conditions for this variable.

Table 7.8. Internal conditioning of yes/no questions for the HF_2 cohort

	TU			ECQ			INV			INT			
Corrected mean			x			0.505			0.046			0.450	
Overall %			0%			50%			5%			45%	
Total N			0			55/109			5/109			49/109	
	FW	%	N	FW	%	N	FW	%	N	FW	%	N	Tot N
Subject type													
2P				[.41]	41%	21	[ ]	10%	5	[.54]	49%	25	51
Others				[.58]	59%	34	K.O.	0%	0	[.47]	41%	24	58
Verb frequency													
High				[.50]	51%	38	[.48]	4%	3	[.50]	45%	34	75
Low				[.50]	50%	17	[.55]	6%	2	[.49]	44%	15	34
Style													
Casual				[.49]	48%	30	[.32]	2%	1	[.55]	50%	31	62
Careful				[.52]	53%	23	[.75]	9%	4	[.43]	37%	16	43

7.5.4. Extralinguistic factors

It seems clear from the internal conditioning above that there are significant differences between cohorts, but Fisher's tests were performed nonetheless for confirmation's sake. Since this variable is not binary (as was the case for the other three), these tests are performed one at a time by pitting each variant against all others.

Table 7.9. Cross-tabulation (2x2 contingency table) of origins and yes/no question variants

***p* < 0.0001**	**TU**		**Others**		**Total N**
LF	74	22%	267	78%	341
HF	0	0%	181	100%	181
Total	**74**	**14%**	**448**	**86%**	**522**
***p* = 0.0366**	**ECQ**		**Others**		**Total N**
LF	115	34%	226	66%	341
HF	78	43%	103	57%	181
Total	**193**	**37%**	**329**	**63%**	**522**
***p* = 0.0929**	**INV**		**Others**		**Total N**
LF	27	8%	314	92%	341
HF	7	4%	174	96%	181
Total	**34**	**7%**	**488**	**93%**	**522**
***p* = 0.0004**	**INT**		**Others**		**Total N**
LF	125	37%	216	63%	341
HF	96	53%	85	47%	181
Total	**221**	**42%**	**301**	**58%**	**522**

Significance calculated using Fisher's two-tailed exact test.

In this case, origins clearly have an important (and quite predictable) effect on variant selection: the LF-origin cohorts significantly favour *-tu* and *est-ce que*, whereas the HF-origin cohorts favour intonation.

Table 7.10. Cross-tabulation (2x2 contingency table) of generation and yes/no question variants

***p* = 0.0420**	**TU**		**Others**		**Total N**
G_1	51	15%	251	85%	302
G_2	23	13%	197	87%	220
Total	**74**	**14%**	**448**	**86%**	**522**
***p* = 0.0045**	**ECQ**		**Others**		**Total N**
G_1	96	28%	206	72%	302
G_2	97	54%	123	46%	220
Total	**193**	**37%**	**329**	**63%**	**522**
***p* = 0.0715**	**INV**		**Others**		**Total N**
G_1	25	7%	277	93%	302
G_2	9	5%	211	95%	220
Total	**34**	**7%**	**488**	**93%**	**522**
***p* = 0.7204**	**INT**		**Others**		**Total N**
G_1	130	38%	172	50%	302
G_2	91	50%	129	47%	220
Total	**221**	**42%**	**301**	**58%**	**522**

Significance calculated using Fisher's two-tailed exact test.

As for generation, *-tu* (a variant associated only with LF-origin cohorts) is associated with G_1 speakers, whereas *est-ce que* (a variant used frequently by all cohorts) is strongly associated with

G_2 speakers. The difference between generations here can be interpreted as well-aligned with the findings from other variants, in the sense that G_1 cohorts maintain distinctions between them, whereas G_2 speakers are generally cohesive – though in this case, the variant they have selected to converge on is quite unexpected from benchmark data.

The extralinguistic conditioning for LF_1 reveals that the significant social effects seem to correspond to the style effects seen for these variants in other studies of LF (Table 7.11).

Table 7.11. Bivariate analysis of extralinguistic effects on yes/no question variant selection for the LF_1 cohort

	TU		ECQ		INV		INT		Tot N
	22%	51	32%	73	10%	23	36%	83	230
Exposure to school Fr.	***p* = 0.036**		*p* = 0.294		***p* = 0.005**		***p* = 0.020**		
Higher	13%	8	37%	23	2%	1	48%	30	62
Lower	26%	43	30%	50	13%	22	32%	53	168
Individual contact	***p* = 0.008**		*p* = 0.477		*p* = 0.155		***p* = 0.014**		
Higher	39%	17	36%	16	5%	2	21%	9	44
Lower	18%	34	31%	57	11%	21	40%	74	186
%Life in Victoria	*p* = 0.186		***p* = 0.022**		*p* = 0.106		***p* = 0.013**		
High	15%	6	27%	11	17%	7	42%	17	41
Moderate	27%	31	40%	46	6%	7	27%	31	115
Low	19%	14	22%	16	12%	9	47%	35	74
Socioeconomic class	*p* = 0.426		***p* = 0.021**		*p* = 0.469		***p* = 0.016**		
Higher	24%	37	37%	57	.60	50	31%	35	156
Lower	19%	14	22%	16	.44	46	47%	48	74

Light gray denotes significant favouring effects.

For example, higher socioeconomic class favours *est-ce que* (a variant associated with hyperformal style), and lower exposure to school French favours *-tu* and inversion (both informal LF variants), but favours intonation (associated with formal style in benchmark LF, also the prototypical HF variant). In other words, there are clear parallels between the extralinguistic conditioning for LF_1 speakers and the style constraints attested for LF_1 (as well as benchmark LF). It is interesting to note that there is no effect of exposure to school French on the use of *est-ce que* for this cohort – indeed, it is the *only* variant for which school French has had no effect whatsoever. I had suggested that LF_1 speakers might be reorganising their question-formation system to use intonation as the 'default' variant rather than *-tu*, and that the rate of *est-ce que* may have been driven up by its increased functional load as the only 'formal' variant in LF_1. The extralinguistic factors do not immediately seem to bear this out – except to suggest that the increase in *est-ce que* may be instigated by speakers with higher socioeconomic status. In other words, the main condition on *est-ce que* for LF_1 would appear to be functional (speech style) rather than extralinguistic, though in this case its function does have some social meaning. This suggests that LF_1 speakers have not shifted away from benchmark norms *entirely*, though their weakened use of *-tu* and increased use of *est-ce que* reflects the social priorities and attitudes within the Victoria milieu (i.e., to avoid 'bad' French).

This conclusion is supported by the effect of proportion of life lived in Victoria. This may appear contradictory at first glance, since both the high and low categories pattern alike (both significantly disfavour *est-ce que*, at 27% and 22%, respectively, as compared to 32%), but the 'moderate' category has contrasting behaviour (significantly favouring *est-ce que*, at 40%). However, this is due to an interaction with the 'individual contact' factor group in this cohort.

All LF_1 individuals with high and low %Life also have low contact with individuals in the broader community, and most are low-contact with infrastructure as well (N = 4/6). In contrast, the 'moderate' category contains individuals with both high and low contact in these two measures. I interpret this to mean that length of time residing in Victoria is not as important as the regular interactions one has over the course of one's residence. In other words, variant selection is affected by the degree to which LF_1 speakers interact with the broader community – lower contact favours intonation, and higher contact favours *-tu*. The retention of *-tu* thus appears to be down to speakers who use French regularly, whereas those who use it less often are the ones susceptible to the shift towards intonation as the default variant.

Table 7.12 shows that the social effects uncovered for LF_2 are not the same as those for LF_1. For instance, use of the intonation variant is favoured by LF_2 individuals who use mostly French in the home and at work, or are relatively unrestricted, or have higher contact with infrastructure – all of which imply a higher degree of participation in the community or more French usage, in terms of frequency or in terms of contextual domains. In contrast, speakers with less occasion to use French (e.g. those who use English some or most of the time in a range of contexts) are the ones favouring *est-ce que*. I interpret this as follows: LF_1 speakers use more *est-ce que* than expected, which has perhaps made it more 'noticeable' to LF_2 speakers, but the LF_2 individuals who regularly use French are (subconsciously) aware that *est-ce que* does not function as the default in benchmark LF, and they instead align themselves with the LF_1 'default' variant, intonation. The same explanation applied to the use of *-tu* in LF_1, which was also associated with the speakers who use French more regularly. This connection is simply expressed using different variants in LF_1 and in LF_2, indicating that the sociostylistic function of these variants – which presumably equates with their social evaluation – is different from G_1 to G_2.

Table 7.12. Bivariate analysis of extralinguistic effects on yes/no question variant selection for the LF_2 cohort

	TU		ECQ		INV		INT		Tot N
	21%	23	38%	42	4%	4	38%	42	111
Language in home	*n.s.*		***p* = 0.014**		*n.s.*		***p* = 0.016**		
Mostly French	19%	14	29%	22	5%	4	47%	35	75
Both	0%	0	75%	6	0%	0	25%	2	8
Mostly English	32%	9	50%	14	0%	0	18%	5	28
Language at work	*n.s.*		***p* = 0.004**		*n.s.*		***p* = 0.036**		
Mostly French	18%	7	21%	8	8%	3	54%	21	39
Both	0%	0	75%	6	0%	0	25%	2	8
Mostly English	23%	11	49%	23	0%	0	28%	13	47
Restriction	*p* = 0.254		*p* = 0.294		*n.s.*		***p* = 0.029**		
None	17%	8	35%	17	6%	3	42%	20	48
Moderate	17%	6	31%	11	4%	1	49%	17	35
Severe	32%	9	50%	14	0%	0	18%	5	28
Infrastructure contact	*p* = 0.673		*p* = 0.071		*n.s.*		***p* = 0.008**		
Higher	18%	7	26%	10	0%	0	55%	21	38
Lower	22%	16	44%	32	6%	4	29%	21	73

Light gray denotes significant favouring effects.

Turning now to the effects uncovered for the HF_2 cohort, it is noteworthy that there are many social factors conditioning the use of inversion (though the token counts are extremely low), which is an informal variant typically associated with spoken LF (Table 7.13).

Table 7.13. Bivariate analysis of extralinguistic effects on yes/no question variant selection for the HF_2 cohort

	ECQ		INV		INT		Tot N
	51%	55	5%	5	45%	49	109
Language at work	***p* = 0.029**		***p* = 0.012**		*p* = 0.254		
Mostly French	22%	4	17%	3	61%	11	18
Both	50%	2	25%	1	25%	1	4
Mostly English	56%	49	1%	1	43%	37	87
Restriction	*n.s.*		***p* = 0.009**		*p* = 0.404		
None	0%	0	33%	3	67%	6	9
Moderate	55%	27	2%	1	43%	21	49
Severe	55%	28	2%	1	43%	22	51
Individual contact	*p* = 0.912		***p* = 0.074**		*p* = 0.401		
Higher	51%	23	9%	4	40%	18	40
Lower	50%	32	2%	1	48%	31	69
%HF in network	*p* = 0.912		***p* = 0.074**		*p* = 0.401		
Higher	50%	32	2%	1	48%	31	64
Lower	51%	23	9%	4	40%	18	45
Socioeconomic class	*p* = 0.653		***p* = 0.081**		*p* = 0.793		
Higher	48%	22	9%	4	44%	20	46
Lower	52%	33	2%	1	46%	29	63

Light gray denotes significant favouring effects.

Inversion shows much of the same conditioning that was attested for intonation in the LF_2 cohort: it is favoured by individuals who use French at least some of the time, by those who are unrestricted, and by those who have higher contact with individuals. Furthermore, use of

inversion is favoured by individuals with lower exposure to HF in their network. Alongside the results for LF_2, I interpret this as evidence for the fact that HF_2 rates of inversion are boosted through contact with LF-origin speakers. The other significant extralinguistic effects are aligned with the LF_2 cohort, thus confirming the cohesion between their underlying patterns of variant selection.

7.6. Summary and interim conclusions

What is clear from this variable is that the G_2 speakers in Victoria are not pulling patterns out of thin air. The presence of implicit usage norms that identify them as LF- of HF-origin speakers (which was not discernible for any of the other variables) clearly indicates that their exposure to HF from parents (or lack thereof) affects their underlying norms.

For both G_1 cohorts, the distribution of variants and the underlying patterns seem to reflect those attested in each of the noncontact benchmarks to some extent, though there is also evidence to suggest a reallocation of stylistic functions for *est-ce que*, *-tu*, and intonation. In LF_1, the diagnostic conditioning for inversion (subject type) is still intact, as well as its style conditioning, which contrasts with that of HF. The style conditioning for *est-ce que* is in place. The style conditioning for *-tu* in LF_1 is not as expected, but this appears to be a consequence of the topic-based style coding protocol. In fact, *-tu* in LF_1 is strongly associated with emotional speech modes such as complaining. Only the intonation variant seems to genuinely contravene the benchmark expectations, since it is favoured in informal, rather than formal, contexts. I suggested that this reflects a shift away from *-tu* as the default variant – a mantle which is taken up by intonation – and that *est-ce que* rates are inflated as a result of this shift, as it takes on more of the functional load of 'formality' previously shared with intonation. In the HF_1 cohort, there is

conditioning that reflects the division of labour between *est-ce que* and intonation in the HF system: intonation is the 'default' form, whereas *est-ce que* is for careful speech.

As for the G_2 cohorts, there are indications that they have made inroads towards a shared, G_2 pattern. Both G_2 cohorts seem to use *est-ce que* as the 'default' variant for question formation – a conclusion supported by multivariate analysis of the internal conditioning. I argue that this cohesive strategy is actually underpinned by the restriction of *-tu* (both in terms of rate, and in terms of stylistic context) by LF_1 speakers. *Est-ce que* is far more frequent in LF_1 speech than in benchmark LF, which I argue is due to its function as the sole 'careful speech' variant in LF_1 – intonation having been reanalysed as the 'default' variant in order to avoid using *-tu*. The LF_2 cohort has the same avoidance of *-tu* (either because they are exposed to it at a lower rate, or for sociostylistic/attitudinal reasons, cf. the extralinguistic and stylistic conditions discussed above), but they lack any style constraint which would block *est-ce que* from acting as the default at LF_1, which leaves it open to take on the role of 'default' in G_2 speech.

Furthermore, the choice of variant is consistent with an effect of the school context, since *est-ce que* is explicitly endorsed by the pedagogical enterprise. Supporting the idea that this increase in *est-ce que* across the board might be due to the school is Elsig's (2009) finding that *est-ce que* rates increased from the 19th to the 20th century (though were still confined to hyperformal contexts). He interprets this increase as being due to a difference in genre between the 19th century and 20th century corpora (Elsig 2009, 122), but I wonder whether it might not be due to the differences in education level of these samples – especially since *est-ce que* is used in only the most formal of utterances in that data. It might be no accident that a variant that is heavily promoted in pedagogical literature (and one that is not subject to implicit usage constraints that narrowly restrict it to certain linguistic contexts, as is the case for inversion – the other

prescriptively-acceptable variant) is the one that G_1 speakers use more of, and the one that G_2 speakers converge on. I had suggested that LF_1 speakers might be motivated to avoid *-tu* given its emblematic association with LF speech and the relatively negative attitudes towards LF expressed by some of my participants. But it might also be the case that the shifts in stylistic treatment of the variants were motivated by a conscious increase to *est-ce que* – perhaps because of its normative appeal. At any rate, this variable unscores the clear connection between salience (itself caused by sociostylistic evaluation of certain variants) and linguistic behaviour in Victoria.

Chapter 8. Finding the big picture: Summary and general discussion

8.1. Review of study aims and methods

The primary aim of this dissertation was to establish whether G_2 speakers in Victoria (a community where French is in the strong minority and is thus contextually restricted) share implicit norms for speech. In many previous studies of communities where two dialects come into contact, a frequently recurring outcome of such contact is the wholesale adoption of locally predominant forms by children, but not by adults – at least in scenarios where certain social conditions are met (e.g., a local dialect is already well-established, cohesive social groups have formed, and community-level interaction is regular and sustained).

The Victoria francophone community instantiates a test of the outcomes of a contact scenario that occurs under different social conditions. Firstly, though LF-origin speakers from outside Victoria are numerous, they are also individually quite transient. As such, no local dialect can be said to be truly 'established': the Victoria francophone community has been maintained over several decades – but not by establishment of sequential generations of speakers. Instead, cycles of renewal feed new individuals into the community, who replace the losses sustained by attrition or out-migration. Secondly, G_2 speakers can be said to have formed cohesive social groups to a large extent, but not necessarily *French-speaking* social groups. In general, the primary language of interaction in these groups tends to be English, not French – even though the individuals in these social groups have the option of speaking French and all understand it perfectly. Finally, community-level interaction occurs only sporadically – though some francophone individuals engage with other francophone individuals and francophone infrastructure far more than others.

Despite these important differences from the social contexts studied in previous work on dialect contact, impressionistic assessments provided by the participants seemed to indicate that G_2 speakers had achieved linguistic cohesion nonetheless. Many individuals identified G_2 speakers as having a characteristically 'Victoria' quality to their French – though opinions varied on what exactly this might mean ('standard', 'mixed', '*Brodeurien*', and so on). Thus, the corollary to the question of whether G_2 speakers share usage norms is the question of how they came to do so and which of the French dialects in their environment might have had more or less influence over their speech. To this point, I hypothesised that G_2 cohesion could be attributed to one of three models of linguistic praxis provided by the broader community: LF-oriented usage, HF-oriented usage, or the school French outlined in grammar books and possibly even modelled by teachers in the school. I argued at the outset that each of these models stood an equally good chance of being adopted by G_2 speakers: LF usage because the community is populated primarily of LF speakers (cf. the determinism stance espoused by Trudgill), HF usage because of its prestigious status and significant presence in the major francophone domains of the community (e.g. the school), and school French because of the important role of the francophone school as a context of French transmission and maintenance, and as a community hub.

In order to assess whether G_2 speakers were aligned with any of these three models, I made use of the comparative variationist method, within the framework of Variation Theory. These methods allowed me to objectively compare underlying patterns for variables that behave differently in each of the three putative source varieties, thereby elucidating which of these models has had the greatest influence on G_2 speech. Comparisons were effected by assessing the rates of the variants of interest, but also – crucially – by statistically modeling the underlying conditioning of the variability. This underlying conditioning can be compared across varieties

using three metrics: statistical significance of the factor groups involved in variability, the constraint hierarchy or direction of effect, as well as the range or magnitude of effect. The focus in the analyses of this dissertation was on variables that would be clearly *diagnostic* of G_2 alignment with HF, LF, or school French, as well as on linguistic and extralinguistic factor groups for each of these variables. These factor groups constitute hypotheses about variant selection, which are then substantiated or refuted through statistical analysis.

8.2. Main findings

The findings have already been extensively summarized in each of the preceding chapters, so I turn now to a more global or wholistic evaluation of some of the main trends uncovered for this community, before tying the findings to questions raised within the literature review. Each of the variables that was included in this study added pieces to the puzzle of G_2 Victoria French.

Assibilation tested whether G_2 speakers were more aligned with HF or LF phonology. Since many studies of dialect contact use variables with similar parameters (near-categorical – or categorical enough to be diagnostic; situated at the phonetic/phonological level of the grammar; etc.), assibilation allowed for some comparability with previous research. Results showed that most G_2 speakers converge on LF-oriented norms, even as G_1 speakers maintain phonological differences, which aligns with the general conclusion that G_2 speakers tend to converge on shared tendencies in situations of dialect contact.

It was observed that the Victoria francophone community believes G_2 speech to be more 'standard' – an unqualified assessment, but one which is directly testable with the methods of comparative variationist analysis. *Avoir/être* tested whether G_2 speakers would conform to a salient and straightforward rule of school French or whether they would instead adopt the implicit patterns associated with usage, which – for this particular variable – are comparable in

HF and LF. Results suggested that usage norms are in place at G_2, though they may have been reanalysed slightly to permit less variability (i.e., *avoir* is categorically avoided in verbs that simply disfavour it in the usage norm, like *arriver* or *partir*, and used at far higher rates than expected in verbs that favour it in the usage norm, like *passer*). It was also apparent from the extralinguistic analysis that sociostylistic factors such as exposure to school French or speech style contribute to the avoidance of *avoir* across the board, though the change in progress away from *avoir* usage that characterises LF communities elsewhere (which is thought to stem from an increase in widespread availability of formal education) may also have played an important role in greatly reducing *avoir* usage overall. This variable provided confirmation that usage-based linguistic or internal *conditioning* could be in place even as external factors relating to negative evaluation or stigma of the variant in question might affect its *rate*.

The future temporal reference variable allowed for a three-way competition between HF usage, LF usage, and school-French prescription, since each of these is attested to have different internal patterns. The findings for this variable verified, in the first instance, that G_1 speakers use distinct usage patterns and that these patterns do not align with prescriptive injunction, which in turn confirmed that G_2 speakers are provided with three competing models to choose from in shaping their own linguistic behaviour. As it turns out, G_2 speakers align with neither LF_1 nor HF_1, instead appearing to conform to a change in progress underway in other parts of the LF-speaking world, where younger generations increasingly gravitate towards the periphrastic variant. The G_2 cohorts were so categorical in their tendency to use the periphrastic future that an assessment of their underlying conditioning was not possible, but what little evidence was uncovered – notably, that speakers who used the synthetic future mostly confined its use to quasi-fixed expressions like *ça sera* – served to support the change in progress explanation for the low rates of synthetic

future at G_2. This variable also shed some light on the connection between stigma and salience, and the importance of these factors in shaping linguistic behaviour. The benchmark usage norms for both HF and LF assign a higher status to the synthetic future than to the periphrastic future, but neither variant is explicitly condemned by grammar books or stigmatised in usage and might therefore be less immediately salient. In my data, both G_1 cohorts were confirmed to have some association of status with the synthetic future, as is the case in benchmark data: in LF_1, increased exposure to school French correlated with increased used of the synthetic future, and the synthetic future was used categorically in careful style contexts by the HF_1 cohort. However, potential positive associations with the synthetic future have done little to encourage its use at G_2, again suggesting that the younger cohorts both simply conform to broader LF trends.

The final variable of this study, yes/no questions, is also the most complex – in terms of analytical legwork, its underlying patterning, and its contribution to multiple questions. For instance, whereas assibilation, *avoir/être*, and future temporal reference were all unable to speak to whether G_2 speakers shared any underlying conditioning (simply because G_2 speakers exhibited rather categorical tendencies for each of these variables), yes/no questions provided the first insights into G_2 *patterns*. Although the categoricity of G_2 linguistic behaviour for the first three variables can, in fact, be equated with G_2 cohesion since there is compelling evidence (particularly from future temporal reference) that G_1 speakers are *not* categorical across the board, this categoriciy prevented any analysis of their underlying patterns. The conclusion of G_2 cohesion is supported by the results for yes/no questions, which show that G_1 speakers maintain distinct underlying systems, but that G_2 speakers behave alike in many respects. Although their rates of each variant are not identical (LF_2 speakers use *-tu*, whereas HF_2 speakers do not), there is evidence to suggest that they have converged on shared underlying conditions for variant

selection – most notably as concerns the variant *est-ce que*, which has completely lost the stylistic associations that it maintained at G_1.

My proposed explanation for this sociostylistic shift at G_2 is rooted in the behaviour of LF_1 – and to a lesser extent, HF_1 – both of which are unusual relative to their respective benchmarks. It seems that LF_1 speakers have reduced use of *-tu* relative to the benchmark, likely because of its strong associations with informal LF vernacular (associations which were confirmed for LF_1 through analysis of sociostylistic factors). They have also increased their rate of intonation and reversed a benchmark constraint on this variant such that it is now favoured with *in*formal speech contexts rather than formal ones. Finally, their rate of *est-ce que* is *vastly* inflated relative to benchmark data, despite maintaining the stylistic constraint that favours its use in more formal speech contexts.

I interpreted these findings for G_1 and G_2 as connected: *-tu* is reduced, intonation acts to fill the space left behind within the 'informal' sphere, and *est-ce que* shoulders the burden of formal contexts alone – the inflated rate here being at least partly due to the fact that LF_1 speakers have more formal tokens than informal ones for this particular variable. This shift at LF_1 would then have effects at G_2, being that LF-origin speakers form the bulk of G_2 input at the level of the community. Both G_2 cohorts show a strong preference for *est-ce que*, but both also lack the stylistic conditioning for this variant that was present at LF_1. This suggests yet another reanalysis of this variant's function: G_2 speakers use *est-ce que* as the 'default' variant – a conclusion supported by their internal and external conditions on variant selection. The role of HF_1 in this paradigm appears to be minor (reflecting their much smaller presence at the broader community level), but their influence is clear in the linguistic behaviour of HF_2: like their HF_1 counterparts, they *never* use *-tu*.

In other words, G_2 speakers of both cohorts are susceptible to effects from the sociostylistic reanalysis (and subsequent boosting of *est-ce que*) that began with LF_1 speakers, yielding cohesive and shared underlying patterns for *est-ce que*, but differences between HF_2 and LF_2 (in terms of *-tu* use) reflect the influence of HF_1 on the former cohort. Since, presumably, HF_2 speakers would receive *some* exposure to *-tu* from their LF_2 peers and from LF_1 adults in their linguistic environment, this finding suggests that exposure to usage within the home takes precedence over usage out in the broader community. Alternatively, this could be interpreted as a strong constraint against adopting any variants that carry some stigma. There is nothing motivating HF_2 speakers to adopt *-tu* – in contrast to LF_2 speakers, whose exposure to *-tu* takes place in the home as well as out in the broader community, and for whom an emblematically LF variant might carry some covert prestige.

In sum, a major advantage of using multiple variables to answer my research questions is that each variable adds to our understanding of linguistic behaviour in Victoria's francophone community, which in turn contributes to questions raised in the broader literature on dialect contact. Globally speaking, the most striking finding of this study is the fact that G_2 speakers have clearly achieved cohesive norms for each of the variables studied here (with but a handful of exceptions, summarized above). Where G_2 speakers (subtly) differed from each other, they tended to align themselves with the G_1 model that was provided in the home; for instance, as far as the usage of the emblematic LF variant *-tu*, or for the three HF_2 speakers who do not have phonologized assibilation.

8.3. Connections to the dialect contact literature

In the introductory material, I had drawn a parallel between the Victoria situation and situations of new-dialect formation or koineisation, in that Victoria does not have the same historical

settlement pattern as many other minority francophone communities in Canada, which gives it the appearance of 'unsettled' linguistic territory. Trudgill (2004) has claimed that G_2 speakers in this type of scenario display 'chaotic' variability as they try to make sense of the multiple models in their environment, with many G_2 speakers using idiosyncratic ('mixed' or 'hybrid') linguistic forms or otherwise innovating relative to the G_1 inputs. He also suggests that G_2 speakers in this type of environment will be more likely to speak like their parents than their peers. Although the Victoria situation is not fully comparable to the colonial New Zealand situation, there is some evidence to suggest that parental norms may play a greater role in Victoria than they would in other, more stable, dialect contact situations, given the small deviations from LF behaviour observed for the HF_2 cohort (e.g. in their avoidance of *-tu* and the incomplete convergence on assibilation). On the other hand, the complete absence of 'chaos' in the speech of G_2 Victoria francophones suggests that Victoria might not be as linguistically unmoored from the broader LF community in Canada as it might initially appear. Perhaps LF acts as a supralocal norm within the Canadian context (though confirming this would require some additional investigation beyond the scope of this dissertation), or it might simply be the case that the overwhelming majority of LF-origin individuals in the Victoria context – despite their general mobility – has provided consistent enough input to G_2 speakers for them to acquire cohesive, largely LF-oriented norms.

Some of the processes thought to be fundamental to new-dialect formation and koineisation (e.g. 'simplification' – the reduction in number of grammatical categories, regularisation of morphology, and so on; cf. Kerswill & Williams, [2000, 67]) are unattested in Victoria data. For example, while assibilation is categorical (and therefore 'regular'), it instantiates a contrast between assibilation productions in one environment and unassibilated productions in other environments, which is maintained across the Victoria population. The evidence from *avoir/être*

also speaks to this directly, since Victoria speakers are *maintaining* the contrast between *avoir* and *être* rather than simply conforming to only one of the available variants. Incidentally, the findings for *avoir/être* can also inform on the question of whether Victoria francophones are truly comparable to L1 francophones or 'native speakers' (as opposed to 'heritage speakers'). Studies of French Immersion students (Knaus and Nadasdi 2001) find a weakening of this contrast in these speakers (who tend to gravitate towards the majority or 'default' variant, *avoir*, across the board) and that Immersion students have failed to acquire the internal conditioning of variability that is attested for benchmark LF. In contrast, G_2 speakers use both *avoir* and *être*, but also – crucially – use them in the expected contexts according to the implicit LF usage norm. This suggests that they can confidently be classed alongside L1 speakers from less restricted francophone communities, rather than alongside 'heritage speakers' or similar.

On the other hand, other processes associated with koineisation or similar outcomes of long-term dialect contact *do* seem to be present. For instance, some of the linguistic patterns I uncovered for G_2 speakers – such as the increased use of *est-ce que* and the loss of its association with careful speech styles – could be interpreted as reallocation (i.e., when "competing forms are retained but assigned different roles either within the grammar or within the sociolinguistic repertoire" [Dodsworth 2017, 333]). But such sociostylistic shifts regularly occur even in contexts *without* dialect contact, and so it seems specious to conclude on this basis that G_2 French is a koiné or similar – especially since the processes thought to involve changes to *internal* structure (e.g. simplification, focusing, etc.) are unattested. In other words, although Victoria seems 'unsettled', there is little linguistic evidence to suggest that the processes associated with situations of new-dialect formation have taken place in Victoria.

The Victoria situation might instead be more comparable to dialect contact scenarios such as those described by Payne (1980) or Otheguy and Zentella (2012), where newcomers conform – with varying degrees of success depending on age at arrival – to the established local variety. Many of the findings in this study support this assessment. For instance, in comparing the two G_2 cohorts to each other, it is clear that convergence does not proceed bidirectionally. In each case of convergence at the G_2 level, HF-origin speakers conform to LF-origin norms – not the other way around – which clearly demonstrates that predominance of a variety at the local (and possibly supralocal) level is an important factor in deciding who converges to whom.

The primacy of LF usage norms is very evident in the patterns of both LF_2 and HF_2 speakers, wherever conclusive evidence for this effect could be found. For example, both G_2 cohorts conform almost universally to the benchmark LF expectations for assibilation. Where evidence was somewhat less conclusive – that is, where underlying linguistic patterns could not be directly assessed for G_2 speakers due to a lack of variability (e.g. in future temporal reference and, to a lesser extent, *avoir/être*) – it did not contradict or rule out the possibility of benchmark LF alignment. Speaking against this somewhat are the results for yes/no questions, where benchmark LF norms were not perfectly replicated by any of the cohorts. Even LF_1 speakers – who would certainly have acquired benchmark LF patterns given their long stints of residence in majority LF-speaking communities – deviated from the benchmark in both rates and underlying conditioning for the variants involved. However, even though LF_1 patterns do not align one-to-one with benchmark LF, they seem to have had some influence over both cohorts of G_2 speakers – at least as far as treatment of *est-ce que*.

The Victoria results can also speak to the question of whether frequency or attitudinal factors play a greater role in determining the likelihood and direction of convergence. The results

elucidated from each of the four variables strongly suggest that the proportion of one dialect relative to the others in a contact situation is important in determining the direction of convergence. Where G_2 speakers converged, they converged on LF, not on HF, which supports Trudgill's stance on frequency to an extent. However, it also seems clear that social or attitudinal factors (e.g. stigma and associated salience of certain variants) play a role in this community. Much of the sociostylistic conditioning associated with benchmark HF and LF is replicated by G_1 speakers in Victoria, which confirms that their evaluation of the variants at study here broadly corresponds to those from benchmark communities. The synthetic future, for example, is associated with high-status speakers in Quebec and in France – as well as in Victoria. What does not seem clear is whether these evaluations *directly* have any bearing on linguistic behaviour. These associations exist, and may therefore guide linguistic behaviour at G_1, but they do not appear to have been directly transferred to G_2 – with the lack of 'high-status' synthetic future being a prime example. If positive associations with higher status were contributing to convergence in this community, we would expect much higher rates of synthetic future at G_2 (and, arguably, at G_1). Furthermore, while internal or linguistic constraints are evidently being transferred to G_2 (where it is possible to assess this, e.g. for yes/no questions), external or social constraints may not be – as is the case for the sociostylistic conditioning on *est-ce que*.

In a sense, this is quite a remarkable achievement for G_2 speakers. Arguably, subconscious implicit linguistic constraints should be more of a 'challenge' for G_2 speakers to acquire than salient or explicit sociostylistic evaluations of forms – *especially* in an environment that seems so disfavourable to acquisition and maintenance of French, and where negative attitudes towards 'bad' French abound (particularly at G_1). But I would instead posit, somewhat ironically, that these G_1 concerns about 'bad' French might actually be contributing to the *lack* of such concerns

at G_2. If G_1 speakers are (consciously or unconsciously) modeling forms that they believe to be more acceptable (as they clearly do for *est-ce que*, for example), they supply G_2 speakers with a rather 'polished' linguistic input, compared to what might be attested for benchmark communities. However, there is no guarantee that G_2 speakers will necessarily *interpret* this input as 'polished'. Forms with positive sociostylistic associations gain those associations within a social context. Victoria's G_2 speakers are not provided with that same social context, and therefore they might not form the same associations with the same linguistic structures. For example, they might hear 'polished' speech frequently at school, if their teachers use a register that would be considered elevated in a different social context. But within the Victoria context, where school is one of the only environments where G_2 speakers are regularly exposed to French, this 'polished' teacher speech might be perceived as perfectly mundane. Reinforcing this perception is the fact that they are taught from grammar books and pedagogical manuals that present the 'polished' forms as being the only options. The primacy of these forms is then further confirmed by the linguistic behaviour of G_1 speakers, who, for example, use abnormally high rates of forms like *est-ce que* compared to benchmark data – thus adding to the impression that *est-ce que* is a perfectly 'typical' informal question-formation strategy for G_2 speakers to then adopt.

Without diachronic confirmation of the linguistic behaviour of G_1 individuals prior to their arrival in Victoria, it is impossible to say whether the Victoria context itself is directly responsible for the differences between G_1 Victoria French and benchmark French as far as these 'polished' forms (i.e., maybe the G_1 speakers in my sample simply arrived with more *est-ce que* than is typical of their benchmark community). However, it seems reasonable to suggest that G_2 linguistic behaviour is in some way shaped by the overrepresentation of these forms at G_1. It also

seems reasonable to emit the conjecture that G_1 speakers are more acutely aware of which forms are stigmatised and which are prestigious – perhaps as a result of their own linguistic insecurity. These insecurities are likely further magnified by the minority context, where 'losing' one's French is a constant concern. The overrepresentation of 'polished' forms makes sense if one considers that many G_1 speakers believe that language 'quality' ensures its survival – this might be a conscious or unconscious strategy to stave off attrition.

8.4. Final reflections

For all its limitations (e.g. low token counts, data imbalances, sampling challenges), this study has done much to elucidate the linguistic behaviour of Victoria's francophone speakers. It is my hope that the empirical work outlined here will be useful in opposing the persistent belief that contextual restriction necessarily leads to functional restriction, which in turn may relieve some of the linguistic insecurity that characterises minority communities across Canada. In no sense of the word can Victoria francophones be considered to have 'inferior' French. Not only are they fully capable of acquiring implicit usage norms that exist completely below the level of conscious awareness (such as the categorical effect of 2P subjects on inversion), they actually use much higher rates of forms that would be considered polished in other francophone contexts. They have maintained full command over a language that is in the extreme minority in Victoria – a laudable accomplishment which demonstrates that concerns over French 'quality' in minority communities are seriously overblown. These results also underline the success of the community and its infrastructure in supporting G_2 acquisition, and the crucial importance of such support in minority communities. Given that much of G_2 dialect contact – as well as most of their French-language socialization and an important amount of their linguistic development – takes place within the school, it is clear that this institution plays a key role in acting as a sort of ersatz

speech community in and of itself. Through countless cycles of population mobility, attrition, and renewal, the school serves as an unwavering source of linguistic stability for the G_2 speakers that pass through its halls. In addition, francophone youth organisations that create enticing opportunities for G_2 francophones to interact in French outside the classroom, as well as offer programs designed to directly address and assuage linguistic insecurity, play a major role in increasing linguistic confidence at G_2. Any efforts to preserve (or – better yet – expand) such francophone infrastructure are therefore sure to directly contribute to the linguistic heritage of minority francophone communities.

BIBLIOGRAPHY

Agnew, Laurette, François Aubert, Gérald Moreau, Gérald Ricard, Marie Robillard, Antoinette Vaucher, and Jean-Paul Vinay. 1987. *Présence francophone à Victoria, C.B. 1843-1987*. Victoria, BC: Éditions Laplante-Agnew.

Allaire, Gratien. 2014. "La francophonie de l'Ouest : Pérénnité, diversité et rapport à l'autre." In *À l'ouest des Grands Lacs : Communautés francophones et variétés de français dans les prairies et en Colombie-Britannique*, 21–68. Quebec: Presses de l'Université Laval.

Anglejan, Alison d', and G. Richard Tucker. 1973. "Sociolinguistic Correlates of Speech Style in Quebec." In *Language Attitudes: Current Trends and Prospects*, edited by Roger W. Shuy and Ralph W. Fasold, 1–27. Washington, DC: Georgetown University Press.

Antes, Theresa A. 2016. "French Interrogative Structures: A New Pedagogical Norm for the 21st-Century Classroom." *Foreign Language Annals* 49 (3): 596–614.

Armstrong, Nigel, and Tim Pooley. 2010. *Social and Linguistic Change in European French*. Palgrave Macmillan UK.

Armstrong, Nigel. 2001. *Social and Stylistic Variation in Spoken French: A Comparative Approach*. Amsterdam, Philadelphia: John Benjamins.

Ashby, William. 1977. "Interrogative Forms in Parisian French." *Semasia* 4: 35–52.

Auer, Peter, and Frans Hinskens. 1996. "The Convergence and Divergence of Dialects in Europe. New and Not So New Developments in an Old Area." *Sociolinguistica* 10 (1): 1–30.

Auer, Peter, Frans Hinskens, and Paul Kerswill. 2008. *Dialect Change: Convergence and Divergence in European Languages*. Cambridge: Cambridge University Press.

Auger, Julie, and Anne-José Villeneuve. 2017. "Using Comparative Sociolinguistics to Inform European Minority Language Policies: Evidence from Contemporary Picard and Regional French." *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 62 (4): 550–75.

Baylon, Christian, and Paul Fabre. 1973. *Grammaire systématique de la langue française avec des travaux pratiques d'application et leurs corrigés*. Paris: Fernand.

Behnstedt, Peter. 1973. *Viens-tu? Est-ce que tu viens? Tu viens? Formen und Strukturen des direkten Fragesatzes im Fransôsischen*. Tûbingen: Narr.

Beniak, Édouard, and Raymond Mougeon. 1989. "Recherches sociolinguistiques sur la variabilité en français ontarien." In *Le français canadien parlé hors Québec: Aperçu sociolinguistique,* edited by Raymond Mougeon and Édouard Beniak, 69–104. Quebec: Presses de l'Université Laval.

Benmamoun, Elabbas, Silvina Montrul, and Maria Polinsky. 2013. "Heritage Languages and Their Speakers: Opportunities and Challenges for Linguistics." *Theoretical Linguistics* 39 (3–4): 129–81.

Bento, Margaret. 2009. "Une étude sociophonétique des affriquées désonorisées en franco-québécois." *Revue québécoise de linguistique* 26 (1): 13–26.

Berns, Janine Katharina Maria. 2013. "Friction between Phonetics and Phonology: The Status of Affricates." PhD dissertation, Utrecht: LOT.

Bérubé, Daniel, B. May Bernhardt, and Joseph P. Stemberger. 2015. "A Test of Canadian French Phonology: Construction and Extended Use." *Canadian Journal of Speech-Language Pathology and Audiology* 39 (1): 62–110.

Beyssade, Claire. 2006. “La structure de l’information dans les questions: Quelques remarques sur la diversité des formes interrogatives en français.” *LINX* 55: 173–93.

Bigot, Davy. 2008. “ ‘Le point’ sur la norme grammaticale du français québécois oral.” Thèse ou essai doctoral accepté, Montréal: Université du Québec à Montréal.

———. 2011. “De la norme grammaticale du français parlé au Québec.” *Arborescences : Revue d’études françaises*, 1: 1–18.

Blanchet, Philippe. 2016. *Discriminations: Combattre la glottophobie*. Paris: Textuel.

Blondeau, Hélène. 2006. “La trajectoire de l’emploi du futur chez une cohorte de montréalais francophones entre 1971 et 1995.” *Canadian Journal of Applied Linguistics* 9 (2): 73–98.

Blondeau, Hélène, and Michael Friesner. 2014. “Manifestations phonétiques de la dynamique des attributions ethnolinguistiques à Montréal.” *Revue Canadienne de Linguistique* 59 (1): 83–105.

Blondeau, Hélène, and Emmanuelle Labeau. 2016. “La référence temporelle au futur dans les bulletins météo en France et au Québec: Regard variationniste sur l’oral préparé.” *Canadian Journal of Linguistics/Revue canadienne de linguistique* 61 (3): 240–58.

Blondeau, Hélène, and Isabelle Lemée. 2020. “Future Temporal Reference in L2 French Spoken in the Laurentian Regions: Contrasting Naturalistic and Instructed Context” 76 (1): 70–90.

Bouchard, Pierre, and Jacques Maurais. 1999. “La norme et l’école. L’opinion des québécois.” Edited by Conrad Ouellon. *Terminogramme (91-92): La norme du français au Québec. Perspectives pédagogiques*, 91–116.

Bourhis, Richard, and Dominique Lepicq. 1993. "Québécois French and Language Issues in Quebec." In *Trends in Romance Linguistics and Philology*, edited by Rebecca Posner and John N. Green, 345–81. Berlin: Mouton de Gruyter.

Brasseur, A. 2009. "Les marqueurs phonétiques de la perception de l'accent québécois." Master's thesis, Université du Québec à Montréal.

Britain, David. 2012. "Koineization and Cake Baking: Reflections on Methods in Dialect Contact Research." In *Methods in Contemporary Linguistics*, edited by Andrea Ender, Adrian Leemann, and Bernhard Wälchli, 1st edn., 219–38. Berlin/Boston: Walter de Gruyter.

Bullock, Barbara E., and Almeida Jacqueline Toribio. 2004. "Introduction: Convergence as an Emergent Property in Bilingual Speech." *Bilingualism: Language and Cognition* 7 (2): 91–93.

Bybee, Joan L. 2006. "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82 (4): 711–33.

Canac-Marquis, Réjean, and Christian Guilbault. 2014. "Remarques sur le français en Colombie-Britannique et la variété de Maillardville: Examen d'un idiolecte." In *À l'ouest des Grands-Lacs: Communautés francophones et variétés de français dans les prairies et en Colombie-Britannique*, edited by Robert Papen and Sandrine Hallion, 277–301. Quebec: Presses de l'Université Laval.

Canada, Office of the Commissioner of Official Languages, and Consortia Development Group. 2010. *Vitality Indicators for Official Language Minority Communities 3: Three Francophone Communities in Western Canada: The British Columbia Francophone Community.* Ottawa: Office of the Commissioner of Official Languages.

Canale, Michael, Raymond Mougeon, and Monique Bélanger. 1978. "Analogical Leveling of the Auxiliary *Être* in Ontarian French." In *Contemporary Studies in Romance Linguistics*, edited by Margarita Suner, 41–61. Washington, DC: Georgetown University Press.

Candea, Maria, Martine Adda-Decker, and Lori Lamel. 2013. "Recent Evolution of Non-Standard Consonantal Variants in French Broadcast News." In *Interspeech, Aug 2013*, 412–16. Lyon, France.

Centre de transfert pour la réussite éducative du Québec. 2019. *Enseigner la grammaire actuelle: Programme de recherche en littératie*. Québec: CTREQ.

Chambers, J. K. 1992. "Dialect Acquisition." *Language* 68 (4): 673–705.

———. 2003. *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*. Wiley.

Chambers, J. K., and Peter Trudgill. 1998. *Dialectology*. Cambridge University Press.

Chavez, Brigitte, and Camille Bouchard-Coulombe. 2011. "Portrait des minorités de langue officielle au Canada: Les francophones de la Colombie-Britannique." Ottawa: Statistique Canada.

Cheng, Lisa Lai-Shen, and Johan Rooryck. 2000. "Licensing Wh-in-Situ." *Syntax* 3 (1): 1–19.

Chevalier, Jean-Claude, Claire Blanche-Benveniste, Michel Arrivé, and Jean Peytard. 1973. *Grammaire Larousse du français contemporain*. Paris: Larousse.

Clements, George N. 1999. "Affricates as Noncontoured Stops." In *Proceedings of LP '98: Item Order in Language and Speech*, edited by Osamu Fujimura, Brian D. Joseph, and Bohumil Palek, 271–99. Prague: The Karolinum Press.

Comeau, Philip, and Anne-José Villeneuve. 2016. "Future Temporal Reference in French: An Introduction." *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique* 61 (3): 231–39.

Conseil scolaire francophone de la Colombie-Britannique. 2007. “Politique H-800-1 culturelle et communautaire.” Vancouver: Conseil scolaire francophone de la Colombie-Britannique.

Corneau, Caroline. 2000. “An EPG Study of Palatalization in French: Cross-Dialect and Inter-Subject Variation.” *Language Variation and Change* 12 (1): 25–49.

Côté, Marie-Hélène. 2014. “Liaison et assibilation en français laurentien.” In *La liaison: approches contemporaines*, edited by Christiane Soum-Favaro, Annelise Coquillon, and Jean-Pierre Chevrot, 9–32. Peter Lang.

Coveney, Aidan. 1996. *Variability in Spoken French: A Sociolinguistic Study of Interrogation and Negation*. Exeter: Elm Bank.

———. 2002. *Variability in Spoken French: A Sociolinguistic Study of Interrogation and Negation*. Bristol & Portland: Elm Bank.

Cox, Terry. 1998. “Vers une norme pour un cours de phonétique française au Canada.” *The Canadian Modern Language Review* 54 (2): 172–97.

Crystal, David. 2003. *A Dictionary of Linguistics and Phonetics*. 5th edn. Oxford: Blackwell.

Dauzat, Albert. 1943. *Le génie de la langue française*. Paris: Payot.

Déprez, Viviane, Kristen Syrett, and Shigeto Kawahara. 2012. “Interfacing Information and Prosody: French *Wh-in-Situ* Questions.” In *Romance Languages and Linguistic Theory 4*, edited by Irene Franco, Sara Lusini, and Andrés Saab, 135–54. Amsterdam: John Benjamins Publishing Company.

Deshaies, Denise, and Ève Laforge. 1981. “Le futur simple et le futur proche dans le français parlé dans la ville de Québec.” *Langues et Linguistique* 7: 23–37.

Detey, Sylvain. 2010. *Les variétés du français parlé dans l’espace francophone: Ressources pour l’enseignement*. Editions OPHRYS.

Dewaele, Jean-Marc. 1999. “Word Order Variation in Interrogative Structures of Native and Non-Native French.” *ITL - International Journal of Applied Linguistics* 123: 161–80.

Digesto, Salvatore. 2019. “Verum a Fontibus Haurire. A Variationist Analysis of Subjunctive Variability across Space and Time: From Contemporary Italian Back to Latin.” PhD dissertation, Ottawa: University of Ottawa.

Dodsworth, Robin. 2017. “Migration and Dialect Contact.” *Annual Review of Linguistics* 3 (1): 331–46.

Druetta, Ruggero. 2003. “Qu’est-ce tu fais? État d’avancement de la grammaticalisation de *est-ce que*. Deuxième partie.” *Linguae & - Rivista di lingue e culture moderne* 2(1): 21–35.

Dumas, Denis. 2001. “Tendances récentes dans la prononciation du français québécois.” In *French Accents: Phonological and Sociolinguistic Perspectives*, edited by Marie-Anne Hintze, Tim Pooley, and Anne Judge, 240–50. London: Association for French Language Studies, Centre for Information on Language Teaching and Research.

Durand, Jacques. 1993. “La phonologie multidimensionnelle moderne et la description du français.” *Journal of French Language Studies* 3 (2): 197–229.

Elsig, Martin. 2009. *Grammatical Variation Across Space and Time*. John Benjamins Publishing Company.

Elsig, Martin, and Shana Poplack. 2006. “Transplanted Dialects and Language Change: Question Formation in Québec.” *University of Pennsylvania Working Papers in Linguistics* 12 (2): 77–90.

Emirkanian, Louisette, and David Sankoff. 1985. “Le futur ‘simple’ et le futur ‘proche.’ ” In *Les tendances dynamiques du français parlé à Montréal, tome I*, edited by Monique Lemieux and Henriette Cedergren, 189–204. Quebec: Office de la langue française.

Etienne, Corinne, and Kelly Sax. 2009. “Stylistic Variation in French: Bridging the Gap Between Research and Textbooks.” *The Modern Language Journal* 93 (4): 584–606.

Fagyal, Zsuzsanna. 2010. *Accents de banlieue: Aspects prosodiques du français populaire en contact avec les langues de l’immigration*. Paris: L’Harmattan.

Fleischman, Suzanne. 1982. *The Future in Thought and Language. Diachronic Evidence from Romance*. Cambridge: Cambridge University Press.

Fleury, Serge, and Sonia Branca-Rosoff. 2010. “Une expérience de collaboration entre linguiste et spécialiste de TAL : L’exploitation du corpus CFPP 2000 en vue d’un travail sur l’alternance futur simple / futur périphrastique.” *AFLS* 16 (1): 63–98.

Foreman, Annik. 2000. “A Longitudinal Study of Americans in Australia,” In *Proceedings of ALS2k, the 2000 Conference of the Australian Linguistics Society*, edited by Keith Allan and John Henderson, 1-10. Melbourne: Monash University Publishing.

Foulet, Louis. 1921. “Comment ont évolué les formes de l’interrogation.” *Romania* 47: 243–348.

Foulkes, Paul, and Gerry Docherty. 1999. *Urban Voices: Accent Studies in the British Isles*. Routledge. Place?

Fox, Cynthia A. 1991. “A Discriminant Analysis of Yes-No Questions in Quebec French.” *Word* 42 (3): 277–93.

Francard, Michel, Joëlle Lambert, and Françoise Berdal-Masuy. 1993. “L’insécurité linguistique en communauté française de Belgique.”

François, Denise. 1974. *Français parlé: Analyse des unités phoniques et significatives d’un corpus recueilli dans la région parisienne*. Paris: Société d’études linguistiques et anthropologiques de France.

Friesner, Michael. 2010. “Une prononciation «tsipéquement» québécoise?: La diffusion de deux aspects stéréotypés du français canadien.” *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique* 55 (1): 27–53.

Fuga, M. 2002. “Attitude Change within Quebec’s Francophone Population.” Unpublished honours thesis, Montréal: McGill University.

Gaiffe, Félix, Simone Jahan, Ernest Maille, Léon Wagner, Ernest Breuil, and Madeleine Marijon. 1936. *Grammaire Larousse du XXe siècle*. Paris: Librairie Larousse.

Genesee, Fred, and Naomi E. Holobow. 1989. “Change and Stability in Intergroup Perceptions.” *Journal of Language and Social Psychology* 8 (1): 17–38.

Gesner, Edward. 1978. “L’emploi des auxiliaires *avoir et être* dans le parler de la Baie Sainte-Marie, Nouvelle-Écosse.” In *Papers from the 2nd Annual Meeting of the Atlantic Provinces Linguistic Association*, edited by George W. Patterson, 16–22. Halifax: Mount Saint Vincent University.

Gess, Randall, Chantal Lyche, and Trudel Meisenburg. 2012. *Phonological Variation in French: Illustrations from Three Continents*. John Benjamins Publishing.

Giles, Howard, and Patricia Johnson. 1987. “Ethnolinguistic Identity Theory: A Social Psychological Approach to Language Maintenance.” *International Journal of the Sociology of Language* 68: 69–99.

Giles, Howard, Anthony Mulac, James J. Bradac, and Patricia Johnson. 1987. “Speech Accommodation Theory: The First Decade and Beyond.” *Annals of the International Communication Association* 10 (1): 13–48.

Giles, Howard, and Tania Ogay. 2007. “Communication Accommodation Theory.” In *Explaining Communication: Contemporary Theories and Exemplars*, edited by B. B. Whaley and W Samter, 293–310. Mahwah, NJ: Lawrence Erlbaum.

Goldvarb Yosemite: A Multivariate Analysis Application for Macintosh. 2015. University of Toronto and University of Ottawa.

Gougenheim, Georges. 1962. *Système Grammatical de La Langue Française*. Paris: Éditions d’Artrey.

Grevisse, Maurice. 1993. *Le Bon Usage. Grammaire Française*. 13th edn. Paris: Duculot.

Grevisse, Maurice, and André Goosse. 2016. *Le Bon Usage*. 16th edn. De Boeck Supérieur.

Grimm, D. Rick. 2010. “A Real-Time Study of Future Temporal Reference in Spoken Ontarian French.” *University of Pennsylvania Working Papers in Linguistics* 16 (2): 83–92.

Grimm, David Ricky Leigh. 2015. “Grammatical Variation and Change in Spoken Ontario French: The Subjunctive Mood and the Expression of Future Temporal Reference.” Doctoral Dissertation, Toronto: York University.

Grimm, Rick, and Terry Nadasdi. 2011. “The Future of Ontario French.” *Journal of French Language Studies* 21 (2): 173–89.

Gudmestad, Aarnes, Amanda Edmonds, Bryan Donaldson, and Katie Carmichael. 2020. “Near-Native Sociolinguistic Competence in French: Evidence from Variable Future-Time Expression.” *Canadian Journal of Applied Linguistics* 23 (1).

Guilbault, Christian. 2007. “Portrait linguistique des francophones de la Colombie-Britannique.” In *Culture et littérature francophones de la Colombie-Britannique : Du rêve à la réalité*, edited by Guy Poirier, 173–200. Espaces francophones, II. Ottawa: Éditions David.

———. 2010. “Représentations des locuteurs du français en Colombie-Britannique.” In *Les voix du français : À travers l’histoire, l’école, et la presse*, edited by Michaêl Abecassis and Gudrun Ledegren, 41–52. Oxford: Peter Lang.

———. 2012. “Langue française en Colombie-Britannique: Aspects linguistiques et pédagogiques.” In *La francophonie en Colombie-Britannique: Mémoire et fiction*, edited by Guy Poirier, Christian Guilbault, and Jacqueline Viswanathan, 147–72. Ottawa: Éditions David.

Guy, Gregory. 1980. “Variation in the Group and in the Individual: The Case of Final Stop Deletion.” In *Locating Language in Time and Space*, edited by William Labov, 1–36. New York: Academic Press.

———. 1988. “Language and Social Class.” In *Linguistics: The Cambridge Survey*, edited by Frederick J. Newmeyer, 37–63. Cambridge: Cambridge University Press.

———. 2018. “LVC Guidelines for Reporting Quantitative Results.” Unpublished manuscript.

Hall, Tracy A, and Silke Hamann. 2006. “Towards a Typology of Stop Assibilation.” *Linguistics* 44 (6).

Hamlaoui, Fatima. 2011. “On the Role of Phonology and Discourse in Francilian French Wh-Questions.” *Journal of Linguistics* 47 (1): 129–62.

Hazen, Kirk. 2006. “The Family.” In *The Handbook of Language Variation and Change*, edited by J.K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 3rd edn., 500–525. Malden, MA/Oxford: Blackwell Publishing Inc.

Holmes, Janet, and Paul Kerswill. 2008. “Contact Is Not Enough: A Response to Trudgill.” *Language in Society* 37 (2): 273–77.

Jamin, Mikaêl. 2004. " 'Beurs' and accent des cités: A case study of linguistic diffusion in La Courneuve." *Contemporary French and Francophone Studies* 8 (2): 169–76.

Jeanjean, Colette. 1988. "Le futur simple et le futur périphrastique en français parlé." In *Grammaire et histoire de la grammaire: Hommage à la mémoire de Jean Stéfanini*, edited by Claire Blanche-Benveniste, André Chervel, and Maurice Gross, 235–57. Aix-en-Provence: Publications de l'Université de Provence.

Junová, Anna. 2016. "L'affrication en français métropolitain: Enquête sociophonétique à Nice." Mémoire de licence, Brno: Université Masaryk.

———. 2018. "Significations sociales de la prononciation affriquée en français de France: enquête perceptive à Rennes." PhD dissertation, Rennes/Brno: Université Rennes 2/Masarykova universita.

Kailuweit, Rolf. 2011. "Être Ou Avoir En Français Standard Contemporain." In *Peregrinatio in Romania*, edited by Renata Enghels, Machteld Meulleman, and Clara Vanderschueren, 397–420. Ghent: Academic Press.

Kastronic, Laura. 2016. "A Comparative Variationist Approach to Morphosyntactic Variation in Hexgonal and Quebec French." PhD dissertation, Ottawa: University of Ottawa.

Kerswill, Paul. 2013. "Koineization." In *The Handbook of Language Variation and Change*, 519–36. John Wiley & Sons, Ltd.

Kerswill, Paul, and Peter Trudgill. 2005. "The Birth of New Dialects." In *Dialect Change: Convergence and Divergence in European Languages*, edited by Peter Auer, Frans Hinskens, and Paul Kerswill, 196–220. Cambridge: Cambridge University Press.

Kerswill, Paul, and Ann Williams. 2000. "Creating a New Town Koine: Children and Language Change in Milton Keynes." *Language in Society* 29 (1): 65–115.

Kim, Hyonsoon. 2001. "A Phonetically Based Account of Phonological Stop Assibilation." *Phonology* 18: 81–108.

King, Ruth, and Terry Nadasdi. 2001. "How auxiliaries be/have in Acadian French" In *Papers from the 24th Annual Meeting of the Atlantic Provinces Linguistic Association*, edited by Patricia Balcom, Louise Beaulieu, and Gisèle Chevalier, 61–72. Moncton: Université de Moncton.

Kircher, Ruth. 2009. "Language Attitudes in Quebec: A Contemporary Perspective." PhD dissertation, London: University of London.

Knaus, Valerie, and Terry Nadasdi. 2001. "*Être ou ne pas être* in Immersion French." *Canadian Modern Language Review/Revue Canadienne Des Langues Vivantes* 58(2): 287–306.

Labov, William. 1964. "Stages in the Acquisition of Standard English." In *Social Dialects and Language Learning*, edited by Roger Shuy, 77–103. Champaign, IL: National Council of Teachers of English.

———. 1966. *The Social Stratification of English in New York City*. 2nd edn. Cambridge: Cambridge University Press.

———. 1972. *Sociolinguistic Patterns*. University of Pennsylvania Press.

———. 1984. "Field Methods of the Project on Linguistic Change and Variation." In *Language in Use: Readings in Sociolinguistics*, edited by John Baugh and Joel Sherzer. Englewood Cliffs, NJ: Prentice-Hall.

———. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Blackwell.

———. 2001. *Principles of Linguistic Change, Volume 2: Social Factors*. Malden, MA/Oxford: Blackwell Publishers Inc.

Le Page, Robert Brock, and Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

Levine, M. 1990. *The Reconquest of Montreal: Language Policy and Social Change in a Bilingual City*. Philadelphia: Temple University Press.

Maurais, Jacques. 2008. *Les québécois et la norme. L'évaluation par les québécois de leurs usages linguistiques*. Quebec: Office de la langue française.

Milroy, James. 2012. "Sociolinguistics and Ideologies in Language History." In *The Handbook of Historical Sociolinguistics*, edited by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 571–84. Malden, MA: Wiley-Blackwell.

Ministère de l'Éducation de la Colombie-Britannique. 2010. *Français Langue Première M à 12: Programme d'études 2010*. Victoria, BC: Province of British Columbia.

Moag, Rodney. 1977. *Fiji Hindi*. Canberra: Australian National University Press.

Morin, Annick. 2017. "Questioning Particles: A Cross-Linguistic Approach to Quebec French Polar Interrogatives." PhD dissertation, Toronto: University of Toronto.

Mosegaard Hansen, Maj-Britt. 2001. "Syntax in Interaction." *Studies in Language* 25 (3): 463–520.

Mougeon, Raymond, and Édouard Beniak. 1989. *Le français canadien parlé hors Québec, aperçu sociolinguistique*. Quebec: Presses de l'Université Laval.

———. 1991. *Linguistic Consequences of Language Contact and Restriction: The Case of French in Ontario, Canada*. Oxford University Press UK.

Mougeon, Raymond, and Terry Nadasdi. 1998. "Sociolinguistic Discontinuity in Minority Language Communities." *Language* 74 (1): 40–55.

Mougeon, Raymond, and Katherine Rehner. 2019. “Patterns of Sociolinguistic Variation in Teacher Classroom Speech.” *Journal of Sociolinguistics* 23 (2): 163–85.

Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Munro, Murray J, Tracey M Derwing, and James E Flege. 1999. “Canadians in Alabama: A Perceptual Study of Dialect Acquisition in Adults.” *Journal of Phonetics* 27 (4): 385–403.

Murphy, Patrick, Philip Monahan, and Margaret Grant. 2016. “Affrication Patterns and Perceptual Tendencies in Canadian and European French.” In *Actes du Congrès de l'ACL 2016/2016 CLA Conference Proceedings*, edited by Lindsay Hracs, 1-14. Calgary: University of Calgary.

Oakes, Leigh, and Jane Warren. 2007. *Language, Citizenship and Identity in Quebec*. Language and Globalization. Palgrave Macmillan UK.

Ostiguy, Luc, and Claude Tousignant. 2008. *Les Prononciations Du Français Québécois*. 2nd edn. Montréal: Guérin.

Otheguy, Ricardo, and Ana Celia Zentella. 2012. *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford, New York: Oxford University Press.

Otheguy, Ricardo, Ana Celia Zentella, and David Livert. 2007. “Language and Dialect Contact in Spanish in New York: Toward the Formation of a Speech Community.” *Language* 83 (4): 770–802.

Patrick, Peter L. 2004. “The Speech Community.” In *The Handbook of Language Variation and Change*, edited by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 573–97. John Wiley & Sons, Ltd.

Payne, Arvilla. 1980. “Factors Controlling the Acquisition of the Philadelphia Dialect by Out-of-State Children.” In *Locating Language in Time and Space*, edited by William Labov, 143–78. New York: Academic Press.

Péronnet, Louise. 1991. “Système des modalités verbales dans le parler acadien du sud-est du Nouveau-Brunswick.” *Journal of the Atlantic Provinces Linguistic Association* 13: 85–98.

Picone, Michael D., and Albert Valdman. 2005. “La situation du français en louisiane.” In *Le français En Amérique du Nord: État présent*, edited by Albert Valdman, Julie Auger, and Deborah Piston-Hatlen, 143–65. Québec: Presses de l’Université Laval.

Pohl, Georges. 1965. “Observations sur les formes d’interrogation dans la langue parlée et dans la langue écrite non-littéraire.” In *Linguistique et philologie romanes. Actes du Xe Congrès international de linguistique et philologie romanes*, edited by Georges Straka, 501–13. Paris: Klincksieck.

Poplack, Shana. 1978. “Dialect Acquisition among Puerto Rican Bilinguals.” *Language in Society* 7 (1): 89–103.

———. 1989. “The Care and Handling of a Mega-Corpus: The Ottawa-Hull French Project.” In *Language Change and Variation*, edited by Ralph W. Fasold and Deborah Schiffrin, 411–51. Amsterdam: Benjamins.

———. 2015. “Norme prescriptive, norme communautaire et variation diaphasique.” In *Les variations diasystématiques dans les langues romanes et leurs interdépendences. Série*

TraLiRo, edited by Kirsten Jeppesen Kragh and Jan Lindschouw. Strasbourg: Société de linguistique romane.

———. 2018. *Borrowing: Loanwords in the Speech Community and in the Grammar*. Oxford, New York: Oxford University Press.

Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane de Andrade Berlinck, Salvatore Digesto, Dora Lacasse, and Jonathan Steuck. 2018. "Variation and Grammaticalization in Romance: A Cross-Linguistic Study of the Subjunctive." In *Manual of Romance Sociolinguistics*, edited by Wendy Ayres-Bennett and Janice Carruthers, 217–52. Berlin, Boston: De Gruyter.

Poplack, Shana, and Nathalie Dion. 2009. "Prescription vs. Praxis: The Evolution of Future Temporal Reference in French." *Language* 85 (3): 557–87.

Poplack, Shana, Lidia-Gabriela Jarmasz, Nathalie Dion, and Nicole Rosen. 2015. "Searching for Standard French: The Construction and Mining of the *Recueil historique des grammaires du français*." *Journal of Historical Sociolinguistics* 1 (1): 13–55.

Poplack, Shana, Allison Lealess, and Nathalie Dion. 2013. "The Evolving Grammar of the French Subjunctive." *Probus* 25 (1): 139–95.

Poplack, Shana, and Stephen Levey. 2010. "Contact-Induced Grammatical Change: A Cautionary Tale." In *Language and Space: An International Handbook of Linguistic Variation, Volume 1: Theories and Methods*, edited by Peter Auer and Jürgen Erich Schmidt, 391–419. Berlin: Mouton de Gruyter.

Poplack, Shana, and Marjory Meechan. 1998. "Introduction: How Languages Fit Together in Codemixing." *International Journal of Bilingualism* 2 (2): 127–38.

Poplack, Shana, Suzanne Robillard, Nathalie Dion, and John C. Paolillo. 2020. “Revisiting Phonetic Integration in Bilingual Borrowing.” *Language* 96(1): 126-159.

Poplack, Shana, and Anne St-Amand. 2007. “A Real-Time Window on 19th-Century Vernacular French: The *Récits du français québécois d’autrefois*.” *Language in Society* 36 (05).

Poplack, Shana, and Sali A. Tagliamonte. 2001. *African American English in the Diaspora*. Malden, MA: Blackwell.

Poplack, Shana, and Danielle Turpin. 1999. “Does the FUTUR Have a Future in (Canadian) French?” *Probus* 11 (1): 133–64.

Potowski, Kim. 2011. “Intrafamilial Dialect Contact.” In *The Handbook of Hispanic Sociolinguistics*, edited by Manuel Diaz-Campos, 579–97. Malden, MA/Oxford: John Wiley & Sons.

Rea, Béatrice. 2016. “« Je m’ai fait mal quand j’ai tombé » : Questions d’auxiliarité et de réflexivité dans le français parlé à Montréal (1971-2013).” *Scriptum: Revue de l’édition 2014 de VocUM* 1.

———. 2020. “A Comparative Analysis of French Auxiliation, with New Evidence from Montréal.” In *Variation and Change in Gallo-Romance Grammar*, edited by Sam Wolfe and Martin Maiden, 213–40. Oxford: Oxford University Press.

Reinke, Kristin. 2005. *La langue à la télévision québécoise: Aspects sociophonétiques*. Quebec: Gouvernement du Québec.

Roberts, Julie. 1993. “The Acquisition of Variable Rules: T,d Deletion and -Ing Production in Preschool Children.” PhD dissertation, Philadelphia, PA: University of Pennsylvania.

Roberts, Nicholas S. 2012. “Future Temporal Reference in Hexagonal French.” *University of Pennsylvania Working Papers in Linguistics* 18 (2): 97–106.

Robillard, Suzanne. 2014. “Great Aspirations: VOT in Bilingual French.” Conference presentation at CVC VIII, Queen’s University, Kingston, ON.

Rochette, Claude, and Édith Bédard. 1984. *La langue des animateurs de la radio et de la télévision francophones au Québec: Une analyse phonétique*. Quebec: Conseil de la langue française.

Romaine, Suzanne. 1982. *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold.

Roussel, Basile. 2016. “J’ai Therefore Je Suis: A Quantitative Analysis of Auxiliary Alternation in Acadian French.” Poster presentation at NWAV 45, Simon Fraser University and University of Victoria, Vancouver, BC.

Russo, Marijke, and Julie Roberts. 1999. “Linguistic Change in Endangered Dialects: The Case of Alternation between *Avoir and Être* in Vermont French.” *Language Variation and Change* 11 (1): 67–85.

Sabourin, Laura, Jean-Christophe Leclerc, Myriam Lapierre, Michèle Burkholder, and Christie Brien. 2016. “The Language Background Questionnaires in L2 Research: Teasing Apart the Variables.” In *Actes du Congrès de l'ACL 2016/2016 CLA Conference Proceedings*, edited by Lindsay Hracs, 1-15. Calgary: University of Calgary

Sammons, Olivia N., Terry Nadasdi, and Raymond Mougeon. 2015. “ ‘Moving’ through the Past: Thirty Years of *Avoir Été* in Ontario French.” *Journal of French Language Studies* 25 (3): 397–422.

Samuels, Michael L. 1975. *Linguistic Evolution: With Special Reference to English*. Cambridge: Cambridge University Press.

Sankoff, Gillian, and Suzanne Evans Wagner. 2006. “Age-Grading in Retrograde Movement: The Inflected Future in Montreal French.” *University of Pennsylvania Working Papers in Linguistics* 12 (2): 203–16.

Sankoff, Gillian, and Pierrette Thibault. 1977. “L’alternance entre les auxiliaires avoir et être en français parlé à Montréal.” *Langue Française* 34: 81–108.

Sankoff, Gillian, and Suzanne Evans Wagner. 2020. “The Long Tail of Language Change: A Trend and Panel Study of Québécois French Futures.” *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique* 65 (2): 246–75.

Siegel, Jeff. 1985. “Koines and Koineization.” *Language in Society* 14 (3): 357–78.

———. 2010. *Second Dialect Acquisition*. Cambridge: Cambridge University Press.

Smith, Jennifer, Mercedes Durham, and Hazel Richards. 2013. “The Social and Linguistic in the Acquisition of Sociolinguistic Norms: Caregivers, Children, and Variation.” *Linguistics* 51(2): 285–324.

Söll, Ludwig. 1983. “L’interrogation directe dans un corpus de langage enfantin.” In *Études de grammaire française descriptive*, edited by Franz-Josef Hausmann, 45–54. Studies in Descriptive Linguistics 9. Heidelberg: J. Groos.

Sorace, Antonella. 2000. “Gradients in Auxiliary Selection with Intransitive Verbs.” *Language* 76 (4): 859–90.

Spini, Mathilde, and Cyril Trimaille. 2017. “Les significations sociales de la palatalisation/affrication à Marseille : Processus ségrégatifs et changement linguistique.” *Langage et societe* 162 (4): 53–78.

Stelling, Louis E. 2011. “The Effects of Grammatical Proscription on Morphosyntactic Change: Auxiliary Variation in Franco-American French.” *Arborescences*, 1: 1-19.

Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. Cambridge University Press.

———. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. John Wiley & Sons.

Tagliamonte, Sali A., and Sonja Molfenter. 2007. "How'd You Get That Accent?: Acquiring a Second Dialect of the Same Language." *Language in Society* 36 (5): 649–75.

Tailleur, Sandrine. 2013. "The French Wh Interrogative System: Est-Ce Que, Clefting?" PhD dissertation, Toronto: University of Toronto.

Tennant, Jeff. 2012. "Laurentian French Phonology in a Majority Setting Outside Québec: Observations from the PFC Hearst Ontario Study." In *Phonological Variation in French: Illustrations from Three Continents*, 313–39. Amsterdam, Philadelphia: John Benjamins.

Terry, Robert M. 1970. *Contemporary french interrogative structures*. Montreal: Éditions Cosmos.

Thelander, Mats. 1979. *Språkliga variationsmodeller tillämpade på nutida burträsktal = Models of linguistic variation applied to present speech in Burträsk*. Stockholm: Almqvist and Wiksell.

———. 1983. "Provinsialismer i inom- och mellanregionalt perspektiv. I: Språk och tradition. Festskrift till Sven Benson." *Svenska Landsmål* 106: 211–31.

Trimaille, Cyril, Maria Candea, and Iryna Lehka-Lemarchand. 2012. "Existe-t-il une signification sociale stable et univoque de la palatalisation/affrication en français ? Étude sur la perception de variantes non standard." *SHS Web of Conferences 1, 3^{e} Congrès mondial de linguistique française*: 2249–62.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford and New York: Basil Blackwell.

———. 1994. *Dialects*. London and New York: Routledge.

———. 1999. “Dialect Contact, Dialectology and Sociolinguistics.” *Cuadernos de Filología Inglesa* 8: 1–8.

———. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Oxford University Press.

———. 2008. “Colonial Dialect Contact in the History of European Languages: On the Irrelevance of Identity to New-Dialect Formation.” *Language in Society* 37: 241–80.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis, and Margaret Maclagan. 2000. “Determinism in New-Dialect Formation and the Genesis of New Zealand English.” *Journal of Linguistics* 36 (2): 299–318.

Valdman, Albert. 2000. “Comment gérer la variation dans l’enseignement du français langue étrangère aux États-Unis.” *The French Review* 73 (4): 648–66.

Vernet, Marie, and Cyril Trimaille. 2007. “Contribution à l’analyse de la palatalisation en français parlé contemporain.” *Nottingham French Studies* 46 (2): 82–99.

Villeneuve, Anne-José. 2017. “A French Spoken Norm under the Radio-Canada Spotlight: Verbal Negation and Quebec Cultural Elites.” *University of Pennsylvania Working Papers in Linguistics* 23 (2).

Villeneuve, Anne-José, and Philip Comeau. 2016. “Breaking down Temporal Distance in a Continental French Variety: Future Temporal Reference in Vimeu.” *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 61 (3): 314–36.

Vinet, Marie-Thérèse. 2001. *D’un français à l’autre: la syntaxe de la microvariation*. Saint-Laurent, Quebec: Les Editions Fides.

Wagner, Suzanne Evans, and Gillian Sankoff. 2011. “Age Grading in the Montréal French Inflected Future.” *Language Variation and Change* 23 (3): 275–313.

Walker, Douglas C. 1984. *The Pronunciation of Canadian French*. Ottawa: University of Ottawa Press.

Walker, Douglas C., and Réjean Canac-Marquis. 2016. "*French in Alberta: A Speaker from Rivière-La-Paix*." In *Varieties of Spoken French*, edited by Sylvain Detey, Jacques Durand, Bernard Laks, and Chantal Lyche, 279-289. Oxford: Oxford University Press.

Walker, James A. 2012. "Form, Function, and Frequency in Phonological Variation." *Language Variation and Change* 24 (3): 397–415.

Walker, James A., and Miriam Meyerhoff. 2013. *Studies of the Community and the Individual*. Edited by Robert Bayley, Richard Cameron, and Ceil Lucas. Vol. 1. Oxford University Press.

Weinreich, Uriel, William Labov, and Martin Herzog. 1968. "Empirical Foundations for a Theory of Language Change." In *Directions for Historical Linguistics*, 98–188. Austin, TX: University of Texas Press.

Willis, Lauren A. 2000. "Être Ou Ne plus Être : Auxiliary Alternation in Ottawa-Hull French." Thesis, University of Ottawa.

Zimmer, Dagmar. 1994. "Le Futur Simple et Le Futur Périphrastique Dans Le Français Parlé à Montréal." *Langues et Linguistique* 20: 213–26.