JE M'AI FAIT MAL QUAND J'AI TOMBÉ:
A REAL- AND APPARENT-TIME STUDY OF AUXILIARY ALTERNATION IN INTRANSITIVE AND PRONOMINAL VERBS IN SPOKEN MONTRÉAL FRENCH (1971-2016)

DISSERTATION SUBMITTED TO THE BOARD OF EXAMINERS OF THE
FACULTY OF LINGUISTICS, PHILOLOGY & PHONETICS
IN FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF
DOCTOR OF PHILOSOPHY IN COMPARATIVE PHILOLOGY AND GENERAL LINGUISTICS

BÉATRICE REA
LADY MARGARET HALL, UNIVERSITY OF OXFORD
TRINITY TERM 2020

CANDIDATE NUMBER: 632696
WORD COUNT: 98,059

Degree Title: DOCTOR OF PHILOSOPHY IN COMPARATIVE PHILOLOGY AND GENERAL LINGUISTICS

DECLARATION OF AUTHORSHIP

Name: BÉATRICE REA

Candidate number: 632696

College: LADY MARGARET HALL

Supervisors: PROF. MARTIN MAIDEN & PROF. ROSALIND TEMPLE

Title of thesis: *JE M'AI FAIT MAL QUAND J'AI TOMBÉ*: A REAL- AND APPARENT-TIME STUDY OF AUXILIARY ALTERNATION IN INTRANSITIVE AND PRONOMINAL VERBS IN SPOKEN MONTRÉAL FRENCH (1971-2016)

Word count: ____**98,059**_____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*The University Student Handbook* Section 8.7; available at https://www.ox.ac.uk/students/academic/student-handbook). ✓

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at https://www.ox.ac.uk/students/academic/guidance/skills?wssl=1 . ✓

The thesis I am submitting is entirely my own work except where otherwise indicated. ✓

It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution. ✓

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs. ✓

I have clearly indicated the presence of all paraphrased material with appropriate references. ✓

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisors. ✓

I have not copied from the work of any other candidate. ✓

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: http://www.admin.ox.ac.uk/statutes/352-051a.shtml). ✓

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted. ✓

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism. ✓

Candidate's signature: …  …….. Date: ……July 29, 2021..

Acknowledgements/Remerciements

I wish to express my sincerest gratitude, first and foremost, to my thesis supervisors: Prof. Martin Maiden and Mr. John Charles Smith, who played a foundation role in this project, as well as Dr. Ros Temple, who took over from JC on his retirement in 2017, for their essential guidance, much-needed patience, and unfailing kindness throughout the last six years.

I want to thank the Clarendon Fund of the University of Oxford, the Social Sciences and Humanities Research Council of Canada, Lady Margaret Hall, the Canada-UK Foundation, as well as the Faculty of Linguistics, Philology and Phonetics of the University of Oxford for their generous financial support allowing me to complete my doctoral studies.

I also want to to acknowledge my DPhil confirmation examiner Prof. Deborah Cameron for her feedback on an earlier version of this project; Dr. Sam Wolfe for his comments on the published version of Chapter 2; Prof. Wolfgang De Melo for his moral and administrative support; the team of the Central University Research Ethics Committe for facilitating the completion of my fieldwork; and the wonderful staff of Lady Margaret Hall who have been most supportive at every stage of this research project and of my stay in Oxford.

I would also like to extend my thanks to an international community of linguists: Gillian Sankoff, for performing a statistical analysis on my behalf on her Montréal French corpora and kindly sharing her results with me for the purpose of this project; Mathieu Avanzi and André Thibault, for generously sharing their *Français d'ici* auxiliary data with me; Anne-José Villeneuve, for her guidance during the initial stages of this research; Raymond Mougeon, for giving me various tips on how to thematically structure my interviews in order to best elicit auxiliary alternation data; and finally the Sociolinguistics Lab at the University of Ottawa – Shana Poplack, Nathalie Dion, and Basile Roussel – for welcoming me and sharing their insights on auxiliary alternation.

On the technical side, I would also like to acknowledge John Coleman and the Phonetics Lab of the University of Oxford for providing me with recording equipment for my fieldwork in Montréal; my dear friend Dr. Xavier Bach for showing me how to use the transcription software *ELAN*; and Daniel Ezra Johnson for providing me constant practical help with *(Shiny) Rbrul* during the statistical analysis stage of this project.

***

Ce projet n'aurait pu voir le jour sans l'inestimable contribution de mes 56 participants. Les mots me manquent pour suffisamment les remercier d'avoir si gentiment accepté de participer à cet étrange exercice, de m'avoir accueillie dans leurs foyers, de s'être ouverts à moi, et d'avoir été si généreux de leur temps et de leurs mots.

J'ai une pensée toute spéciale pour trois d'entre eux qui ne sont plus parmi nous aujourd'hui : Louise B., Janie L. et Guy N. parce qu'il paraît que l'une des premières choses que l'on tend à oublier d'un être aimé qui nous quitte c'est le son de sa voix. J'utilise ici leur vrai nom, pour la seule et unique fois, dans le but de leur rendre hommage.

Je veux aussi remercier du fond de mon cœur Céline Brault (♥ pour son dévouement), mes amis Tamila Ziani, Ashley Sévigny, Joëlle Shaw, Naomi Bisson, Léticia Villeneuve, Adèle Aubin, Anne-Marie Messier, Monelle Guertin, Simon-Pierre Chevarie-Cossette, Gabrielle Barrieau, ainsi que mon cousin Alain Vadeboncœur qui ont épluché leur liste de contacts afin de trouver des participants pour mon étude.

***

This thesis could not have been completed without the inexhaustible support of my friends and family, who have been especially caring and encouraging in the toughest moments of this looooong journey.

A big thank you to the wonderful friends whom I have met in Oxford and who have truly made my time there unforgettable: Joachim Mowinckel (♥ et dont l'incroyable papi, Jacques Testart, m'a fait découvrir le joli texte de Duteil présenté à la page suivante), Daniel Straulino, Cornelius Probst, Anna Frangou, Ria Kapoor (my Oxford wife), Dom Barker, Sean Kelly, Rudy Bunel, Haya Alfarhan, Smriti Sawkar, Melissa Godin, Kieran Gilfoy, Arthur Asseraf, and Joe Carson-Parker. In the interest of brevity, I have kept the list short and there are so many other names that I could have mentioned, but I trust that you know who you are.

Je veux aussi remercier du fond du cœur mes Québécois d'Oxford, ma deuxième famille, grâce à qui je me suis toujours sentie à la maison et n'ai jamais affronté d'obstacles seule : Manon Deslauriers (♥ ma maman adoptive), Justine Aspireault-Massé, Pascale Gourdeau, Marie-Joëlle Guilbeault, Guillaume Sylvain, Benjamin Mappin-Kasirer, Léo Bureau-Blouin, Léticia et les autres. Vous êtes bien plus nombreux que ce que ces brefs remerciements me permettent de mentionner, mais vous vous reconnaîtrez. Vous n'avez même pas idée à quel point vous m'êtes chers et à quel point je serai toujours reconnaissante de l'affection que vous me portez.

Je voudrais aussi remercier ici les amis qui ont choisi d'adopter une approche plus 'hands-on' et m'ont encadrée dans ma rédaction quand la motivation et la santé n'étaient pas au rendez-vous : Justine et Pascale, Marc-Antoine Fiset, Alicia Cleary-Venables, Sandro-Angelo De Thomasis (pour la relecture du chapitre le plus lourd), Caroline Loranger, Maxine Visotzky, Tamila (♥ pour nos sessions de 'consulting' à la PVM et notre colocation covidienne), et ma cousine Stéphanie Martel (pour les nombreuses relectures de dossiers de bourse). Merci aussi à Éric D'Amours d'amour pour les traductions 'rock' 🤘, les encouragements, ainsi que le soutien gastronomique des derniers milles.

Un gros merci va aussi à mes amis montréalais, qui m'ont été d'un soutien indéfectible, même lorsqu'un océan nous séparait et lors de mon retour à Montréal : Julien Roy-Lavallée, Catherine (pour le third wheeling thérapeuthique ♥) et Charles Frémont, mes princesses que je n'ai pas déjà mentionnées : Emmanuelle Bélanger, Marie-Claire Déry et Bao Chau Phan, ainsi que mes tchieks que je n'ai pas encore nommées : Gabrielle Soucy-Girard et Ariane Perreault.

J'ai aussi une pensée pour Letitia Alexe qui m'accompagne virtuellement en consultation psychologique depuis 2015. Cette thèse a été rédigée en bonne partie en période de confinement covidien et, dans ces conditions aussi peu favorables à la concentration, je n'aurais jamais eu la force de la compléter sans son soutien moral.

Je souhaite aussi remercier *con tutto il cuore* mes trois parents adorés, qui n'ont cessé de me soutenir pendant toutes ces années, qui m'ont toujours encouragée à me surpasser, et à qui je dédie cette thèse de doctorat.

AVOIR ET ÊTRE

Loin des vieux livres de grammaire,
Écoutez comment un beau soir,
Ma mère m'enseigna les mystères
Du verbe être et du verbe avoir.
Parmi mes meilleurs auxiliaires,
Il est deux verbes originaux.
Avoir et Être étaient deux frères
Que j'ai connus dès le berceau.
Bien qu'opposés de caractère,
On pouvait les croire jumeaux,
Tant leur histoire est singulière.
Mais ces deux frères étaient rivaux.
Ce qu'Avoir aurait voulu être
Être voulait toujours l'avoir.
À ne vouloir ni dieu ni maître,
Le verbe Être s'est fait avoir.
Son frère Avoir était en banque
Et faisait un grand numéro,
Alors qu'Être, toujours en manque.
Souffrait beaucoup dans son ego.
Pendant qu'Être apprenait à lire
Et faisait ses humanités,
De son côté sans rien lui dire
Avoir apprenait à compter.
Et il amassait des fortunes
En avoirs, en liquidités,
Pendant qu'Être, un peu dans la lune
S'était laissé déposséder.

Avoir était ostentatoire
Lorsqu'il se montrait généreux,
Être en revanche, et c'est notoire,
Est bien souvent présomptueux.
Avoir voyagé en classe Affaires.
Il met tous ses titres à l'abri.
Alors qu'Être est plus débonnaire,
Il ne gardera rien pour lui.
Sa richesse est tout intérieure,
Ce sont les choses de l'esprit.
Le verbe Être est tout en pudeur,
Et sa noblesse est à ce prix.
Un jour à force de chimères
Pour parvenir à un accord,
Entre verbes ça peut se faire,
Ils conjuguèrent leurs efforts.
Et pour ne pas perdre la face
Au milieu des mots rassemblés,
Ils se sont répartis les tâches
Pour enfin se réconcilier.
Le verbe Avoir a besoin d'Être
Parce qu'être, c'est exister.
Le verbe Être a besoin d'avoirs
Pour enrichir ses bons côtés.
Et de palabres interminables
En arguties alambiquées,
Nos deux frères inséparables
Ont pu être et avoir été.

– Yves Duteil

TABLE OF CONTENTS

Abstract

My thesis investigates auxiliary alternation in spoken Montréal French between *avoir* 'have' and *être* 'be' in the twenty or so intransitive verbs that prescriptively require the latter (*tomber* 'to fall', *partir* 'to leave', *rester* 'to stay', etc.):

(1) *J'ai tombé* (AVOIR) vs *Je suis tombé* (ÊTRE) I fell/have fallen

This phenomenon has been documented in virtually all the French-speaking communities of North America. Sankoff & Thibault (1977, 1980) observed a rate of *avoir* generalization of 34% in the Sankoff-Cedergren Montréal Corpus (1971), and hypothesized that greater exposure to the standard "acts as a brake on the regularization of conjugation through the extension of *avoir*" (1980: 345). Using a trend study, I attempt to determine whether their 'prediction' is confirmed in real time and in apparent time. I also explore the alternation between auxiliary *avoir* and *être* within pronominal forms because they were not studied by Sankoff & Thibault:

(2) *Je m'ai fait mal* (AVOIR) vs *Je me suis fait mal* (ÊTRE) I (have) hurt myself

In 2016, I interviewed 48 native speakers of Montréal French and, with the mixed-effects statistical software *Shiny Rbrul* (Johnson 2017), I test various (socio)linguistic factors on the recorded intransitive and pronominal periphrastic tense tokens in order to determine the main influences on *avoir* selection.

The real-time comparison of my results with those of Sankoff & Thibault (1977, 1980) reveals that the auxiliary alternation observed in intransitive verbs has, overall, significantly decreased in Montréal French. The absence of any age effect in apparent-time suggests that in 2016 the change to *être* has almost reached completion, despite the lingering of a small stable variation in the lowest SPS. My data also show that *avoir* use in the periphrastic tenses of pronominal verbs is highly socially marked. This decline of *avoir* in Montréal French appears to illustrate a (re)alignment with standard French or *dévernacularisation*, thus confirming Sankoff & Thibault's conclusions.

I then triangulate these findings through an analysis of grammaticality judgements performed by my 48 participants and of self-reporting judgements on auxiliary choice from 821 Montréalers of the crowdsourced corpus *Français de nos régions* (Avanzi *et al.* 2016). My 48 speakers were much more willing to accept *avoir* variants, with pronominal verbs even more so than with intransitive ones, than what could have been inferred from their own actual auxiliation patterns. As for the self-reporting judgements, the effects of the level of education of the participants and their gender were highly significant on the choice of auxiliary verb.

By investigating both intransitive and pronominal verbs and by analysing newly collected (production and grammaticality judgement) data, this sociolinguistic study of auxiliary alternation in Montréal French is the first of its kind.

List of Tables

List of Figures

Chapter 1. Introduction: DR & MRS VANDERTRAMP and auxiliary alternation

1.1. Introducing the variable: Sankoff & Thibault (1977) and Thibault & Sankoff (1997)

In Modern French, as in many other Romance and Germanic languages, the periphrastic verb tenses are formed by combining a past participle with the conjugated auxiliary *avoir* 'have'[1,2] or *être* 'be', depending on the verb and the structure of the argument. In prescriptive works, the distinction between verbs that take the auxiliary *avoir* and those that take the auxiliary *être* is straightforward: while all non-pronominal transitive verbs require auxiliary *avoir*, a small subset of intransitive (stative and motion) verbs combine with the auxiliary *être*. The verbs which prescriptively require the *être* auxiliary in French are not normally taught in classrooms to native speakers but have to be learned by heart by second-language learners, since no common semantic features unite them. The mnemonic device for L2 learners, DR & MRS VANDERTRAMP, is a list where each letter stands for the first letter of the intransitive verb needing to be conjugated with *être*, typically *Devenir* ('to become'), *Revenir* ('to come back') & *Monter* ('to go up'), *Rester* ('to stay'), *Sortir* ('to go out'), *Venir* ('to come'), *Aller* ('to go'), *Naître* ('to be born'), *Descendre* ('to go down'), *Entrer* ('to go or come in'), *Rentrer* ('to go back or come back in'), *Tomber* ('to fall'), *Retourner* ('to return'), *Arriver* ('to arrive'), *Mourir* ('to die'), and *Partir* ('to leave'), as well as all their iterative forms.

I first became interested in auxiliaries in 2013, after attending a lecture given by Dr. Ros Temple on linguistic variation and change in the French language. While discussing the different varieties of North American French[3], she listed some typical features of Montréal

[1] For ease of reading, the verbs appearing in French will only be glossed the first time they occur in the text.
[2] All the translations are my own, except where otherwise indicated.
[3] North American French, and Canadian French in particular, is traditionally divided into two major linguistic varieties: Laurentian French and Acadian French (Martineau 2018: 297-301). Laurentian French refers to the collective label that includes the variety of French that is spoken in the St. Lawrence River valley (Québec), those that originate from it through population migrations towards the West (around the Great Lakes, in Western

French, a variety of Laurentian French[4], that had been observed by linguists, among them the use of the auxiliary *avoir* with the twenty-or-so intransitive verbs prescriptively requiring *être*, as illustrated in (1a-b) with *rester* and *partir* in (2a-b).

(1a) *J'**ai** resté*
Lit. 'I have stayed'

(1b) *Je **suis** resté*
Lit. 'I am stayed (I stayed/I have stayed)'

(2a) *J'**ai** parti*
Lit. 'I have left'

(2b) *Je **suis** parti*
Lit. 'I am left (I left/I have left)'

As a native speaker of Montréal French, I was utterly surprised to learn this since I had never been aware that this morphosyntactic feature was part of the variety that I spoke. I had heard a few people use *avoir* with *tomber*, but had simply assumed that this specific instance was very socially marked. My first reaction when faced with this new information was therefore denial: I knew no one in my speech community that spoke in that way. After Ros Temple suggested I read the 1977 paper where the Montréal auxiliary alternation[5] data had come from, I found

Canada, and the American West) and in New England, United States. Acadian French normally includes the varieties spoken in the Atlantic provinces of Canada (New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland) but also in the Îles-de-la-Madeleine, the south coast of Gaspésie, and the Basse Côte-Nord, in the American Madawaska, as well as in Southern Louisiana. Martineau (2018: 300) insists on the fact that due to the extended contact in some regions between Laurentian and Acadian French, many features of one variety can appear in the other and vice versa. Martineau (2018: 300) also warns readers, however, that the collective label *français laurentien* unfortunaly seems to erroneously suggest the existence of an 18th-19th immutable Québécois French. Many researchers have used the term *français québécois* as the collective label for all French varieties whose origin is Québécois French, but Martineau (2018: 300) and Remysen (2021, personal communication) explain that varieties of French that originate from Québécois French but are outside of its political borders can benefit from being classified under a different name, namely *français laurentien,* used in a restricted sense. Both the collective and the restricted meanings of *français laurentien* are reported in Usito (Cajolet-Laganière, Martel, & Masson 2013), the first online dictionary of Standard Québec French (see §2.1.2). In this research, the term 'Laurentian French' will refer to the collective label that includes Québécois French and 'Québécois French' will be used to refer to the variety of French spoken in the province of Québec.

[4] See footnote 3.

[5] Alternation is understood here at the levels of both the speech community and the individual speaker.

that I could not relate to any of the recorded tokens mentioned by the authors, and I was left with two unanswered questions: had I been unaware of the existence of the *avoir* variants because auxiliary alternation is not a variable that is salient to speakers? Or rather because the phenomenon was not as prevalent as it once was in the 1970s?

The paper in question, Sankoff & Thibault (1977), analysed a Montréal corpus (Sankoff *et al.* 1976) of spoken French dating back to 1971. I was therefore encouraged by my MPhil dissertation supervisor, John Charles Smith, to start listening carefully to my twenty-first-century peers and I indeed became slowly attuned to the phenomenon: I noticed that we did indeed generalize[6] *avoir*, nonetheless to, what seemed to me, a much smaller extent than what Sankoff & Thibault had claimed to be the case in 1971. My MPhil dissertation (Rea 2014) would therefore be an update of this survey: in 2013, I conducted a pilot study with 12 native speakers of Montréal French, where I revisited Sankoff & Thibault's seminal 1977 paper[7]. This lay the groundwork for the present DPhil project.

After having analysed the Sankoff-Cedergren corpus of spoken Montréal French from 1971 including sociolinguistic interviews with 120 speakers (±150 hours of recording) and circumscribing the variable context by removing all tokens where *être* and *avoir* did not behave as auxiliary verbs, Sankoff & Thibault (1977) recorded variation in the auxiliary selection of 17 intransitive verbs. In order of probability of appearing with *avoir*, they are: *demeurer* 'to stay/reside'*, rester, changer* 'to change'*, passer* 'to pass/go through'*, déménager* 'to move (houses)'*, tomber, rentrer, sortir, monter, descendre, retourner, partir, arriver, entrer*, *venir*, *revenir*, and *aller*. Their results reveal that women, as well as people with greater exposure to the standard language, speakers of a higher socioeconomic status, older speakers, and speakers who had completed

[6] Like Sankoff & Thibault (1977: 106), I use the terms "generalize" and "generalization" here with a synchronic meaning, i.e. the process whereby an expected auxiliary *être* (with respect to the standard language) is replaced with auxiliary *avoir*; as opposed to the diachronic process whereby, in a language that used to employ two auxiliary verbs, one auxiliary verb is eventually replaced by the other.

[7] An English version of the paper, "The alternation between the auxiliaries *avoir* and *être* in Montréal French", was published in 1980.

higher levels of schooling displayed a more conservative use of the auxiliary, namely greater retention of *être*. Sankoff & Thibault (1977: 107) concluded that greater exposure to the standard significantly inhibits *avoir* use in the periphrastic tenses of intransitive *être*-verbs[8] (i.e. the twenty-or-so verbs prescriptively requiring the auxiliary *être*).

In an unpublished 1997 study, Thibault & Sankoff analysed auxiliary alternation in a Montréal French corpus dating from 1984. Their results showed that the social differentiation was not straightforward: although there was a tendency for the non-standard use of *avoir* to be associated with working-class speakers, this applied differentially to the various verbs, and the social class-based pattern was far from categorical. Moreover, a semantically driven association between auxiliary use and verb meaning was weakly motivated at best. Rather, there seemed to be a lexically-based distribution. Lastly, Thibault & Sankoff (1997) saw little change over time, between 1971 and 1984, and no indication that one competing form was replacing the other (Thibault & Sankoff 1997).

When the present study was almost complete, Sankoff (2019) published a follow-up study focussing on the various trajectories of linguistic change across the lifespan and that was carried out with updated statistical tools on combined 1971, 1984, and 1995 data. In this paper, she reveals that her 1997 preliminary results were not totally accurate. Sankoff (2019: 220) recorded lower *avoir*-selection rates in younger speakers than in older speakers, and this apparent-time interpretation (see §2.5.) led her to conclude that the change towards a diminution of *avoir* was indeed on-going (Sankoff 2019: 221). Moreover, Sankoff (2019: 198) found that patterns of auxiliary alternation were stable throughout an individual's lifespan. Results from this paper will be discussed in more detail in the following chapter and in §4.2.6. and §6.2.1.

[8] They will henceforth be referred to as '*Ê*-verbs'.

1.2. Structure of the thesis

With the creation of a new corpus of spoken Montréal French comprising sociolinguistic interviews with 48 native speakers, this thesis aims to answer the research questions, outlined at the end of Chapter 2, arising from a critical review of the available literature that explores 1) auxiliary selection in French from both the historical and the prescriptive perspectives; 2) split auxiliary selection and auxiliary alternation as Romance phenomena, by studying common *2-aux* 'two auxiliary' systems; 3) the debate surrounding the nature of auxiliary alternation as a sociolinguistic variable; 4) auxiliary alternation in French in particular, by analysing variable auxiliation patterns recorded in both European and North American varieties; and 5) interpretation and methodology challenges to apparent- and real-time analyses. Chapter 3 will then provide a detailed explanation of the sociolinguistic methodology – approved by the Central University Research Ethics Committee[9] (CUREC) of the University of Oxford – employed to 1) constitute a new corpus including production data as well as grammaticality judgement data on auxiliary alternation; 2) operationalize the (socio)linguistic variables; and 3) circumscribe the variable context and transcribe the relevant data; 4) exclude some exceptional tokens from the corpus; 5) subsequently analyse the collected data; and 6) triangulate the results obtained with the production data with grammaticality judgements from my 48 speakers and self-reported judgements from the crowdsourcing survey *Français de nos régions* (Avanzi *et al.* 2016).

Multivariate statistical analyses producing mixed-effects models and distributional analyses of the various (socio)linguistic variables that condition the use of auxiliary *avoir* in both intransitive and pronominal verbs are performed in Chapter 4. These results are compared with older Montréal French data collected in 1971, 1984, and 1995. This chapter also provides an analysis of various avoidance mechanisms, such as the use of *avoir été* 'to have

[9] See ethical approval in Appendix 3A.

been' for *être allé* 'to have gone' and the use of historic present.

Chapter 5 attempts to triangulate the findings of the previous chapter through an analysis of grammaticality judgements performed by the 48 participants in my production study and of a crowdsourced corpus of self-reporting judgements on auxiliary alternation (*Français de nos régions,* Avanzi *et al.* 2016). In Chapter 6, I present a discussion of my findings in the Romance context and of the implications of my results obtained for both intransitive and pronominal auxiliation patterns. I also outline potential directions for future research on the topic of auxiliary alternation. The conclusion of the thesis appears in Chapter 7.

Chapter 2. Review of the literature[10]

This chapter considers the phenomena of auxiliary selection and alternation through a brief overview of French auxiliation from both diachronic and prescriptive perspectives, and then of common auxiliation patterns in Romance, from the simpler to the more complex ones. It discusses theoretical approaches, both syntactic and semantic, put forward in order to synchronically explain split auxiliary selection in French and Romance. The chapter will then summarise the debate surrounding the nature of auxiliary alternation as a sociolinguistic variable. This review of the available literature subsequently provides a description of the Montréal French corpora used in previous studies of auxiliary alternation, and also provides an evaluation of the various variables, both external/social and internal/linguistic, shown to influence *avoir* generalization in North American varieties of French. The chapter will then describe the potential interpretation and methodology challenges of using apparent- and real-time sociolinguistic analyses. Lastly, this chapter outlines the most important gaps in the theoretical approaches and explains why a variationist sociolinguistic framework might yield desirable complementary results. The research questions that stem from this literature review and that underpin this study will be formulated at the very end of the chapter.

2.1. French auxiliation

2.1.1. From a diachronic perspective

Rather than being expressed periphrastically, the various tenses of Latin, a mainly synthetic language, were expressed synthetically, in the active voice, through the addition of desinences to the stem of the verb which indicated its aspect, as well as the number and person of the

[10] An adapted version of this chapter appears in Rea (2020).

subject performing the action. In the passive voice, however, the perfective forms (perfect, pluperfect, and future perfect of the indicative, the perfect and pluperfect of the subjunctive), were formed via a periphrasis comprising auxiliary *esse* 'be' and the past participle (so that the periphrasis containing the present tense of the auxiliary was a present perfective, that with the imperfect the past perfective, etc.). The imperfective forms of the passive were synthetic: a form such as *amor* therefore meant 'I am loved' and *amatus sum* 'I was/have been loved'. Numerous exceptions to this rule concern deponent verbs which were passive in form (conjugated in the perfective with auxiliary *esse* and the past participle) but active in meaning. For example, *secutus sum*, of the verb *sequi*, meant 'I followed' and not 'I was followed'.

The use of *être* with lexical intransitive verbs in French appears to be a construction inherited from the perfect tenses of the Latin deponent verbs (Marchello-Nizia *et al.* 2020: 1495). The synthetic passive morphology of the imperfective was lost, giving rise to the type **morit* 'he dies' - *mortuus est* 'he (has) died' (for Classical Latin *moritur* - *mortuus est*) in formerly deponent verbs. Levitt (1979: 26) and Ledgeway (2011: 454f.) claim that the structural model created by such remnants of intransitive deponent verbs formed the basis for innovations such as **est venutus* (becoming *il est venu* 'he came/has come') or **est intratus* (becoming *il est entré* 'he went in/has gone in').

Fontaine (1888) suggests another hypothesis, namely the influence of pronominal verbs: Old French possessed many more pronominal verbs than Modern French, co-existing alongside their non-reflexive counterparts (Gougenheim 1979: 185; Buridant 2000: §235): *s'en aller* 'to leave' and *aller*, *se mourir* 'to be dying' and *mourir*, *s'en venir* 'to come (here)' and *venir*, as well as *se partir* 'to leave' and *partir*. Buridant (2000: §235) points out that the pronominal structure tended to carry a durative aspect. The use of *être* with intransitives could thus potentially be just a "shortened form" of a former reflexive verb (Fontaine 1888, quoted in Levitt 1979: 26). Though *s'en aller* 'to leave' and *se mourir* 'to be dying' have survived, most doublets have disappeared from Modern French. However, reflexive forms such as *s'en*

(re)venir are still found[11] in spoken Québécois French, as reported in dictionary *Usito* (Cajolet-Laganière, Martel, & Masson 2013), the first online dictionary of Standard Québec French (see §2.1.2). This verb is illustrated in (3).

(3) *Je m'en suis (re)venue à pied / Je me suis t'en-(re)venue à pied*
'I came (back) here on foot.' (My example)

As early as Old French, the distribution of auxiliaries was similar to that attested in Modern French: transitive and most intransitive verbs formed their periphrastic tenses with auxiliary *avoir*, while passive forms and pronominal verbs did so with *estre* (*être*) (Buridant 2000: §225, §299; Marchello-Nizia *et al.* 2020: 836). The few (change of location or of state) intransitive verbs that combined with *être* comprised *aler* 'to go'*, venir, entrer, issir* 'to leave, to go out'*, partir, devenir, cheoir* 'to fall'*, naistre* 'to be born'*,* and *morir* 'to die' (Moignet 1973: 183; Marchello-Nizia *et al.* 2020: 836). However, Marchello-Nizia *et al.* (2020: 836) point out that verbs *aler* and *morir* were found with both auxiliaries.

For Buridant (2000) there are additional parameters to take into consideration. Intransitive verbs were conjugated with *avoir* or *estre* depending on the "mode d'action" '*Aktionsart*': perfective verbs, "portant en eux-mêmes leur finalité et tendant vers un état"[12], normally appeared alongside *estre*, whereas imperfective verbs, which can express an indefinite process, were normally conjugated with *avoir* (Buridant 2000: §299). This opposition can be seen in the following pairs, where the first element appears with *avoir* and the second with *estre*: *dormir* 'to sleep' vs *endormir* 'to fall asleep'*, veillier* 'to stay awake' vs *esveillier* 'to wake up'*,* and *seoir* 'to be sat' vs *aseoir* 'to sit' (Buridant 2000: §299). With *aler* specifically, its derived forms with prefixes *par-, pur-,* and *tres-* (marking the completion/finalisation of the process) and *mes-* (marking its deviation) seem to appear most frequently with auxiliary *estre* (Buridant

[11] Four such tokens were found in the 2016 corpus.
[12] 'carrying their end in themselves and tending towards a state'

2000: §299). Buridant (2000: §299) also points out that another parameter that could often influence auxiliary selection with motion verbs in Old French is whether or not the verb is impersonal (*estre* is always found in the impersonal forms of *aler,* for example), and whether or not the subject is animate (inanimate subjects almost always appear alongside *estre,*[13] while animate subjects, in the vast majority of cases, appear with *estre* if the focus is on the completion/outcome of the process and with *avoir* if the emphasis is on the duration of the process). However, a crucial point to note is that metric factors related to versification in Old French could condition the choice of using either auxiliary for the same structure (Buridant 2000: §299).

Alternation in French auxiliation can also be found in 16th-century texts (Gougenheim 1951: 122; Rideout 2011, 2013; Marchello-Nizia *et al.* 2020: 836) and there is also evidence that in the 17th and 18th centuries (Tailleur 2007; Rideout 2013) certain verbs were often freely conjugated with both auxiliaries interchangeably: among others, the verbs *demeurer*, *descendre*, *entrer, monter, (re)partir, retourner, sortir,* and *tomber* (Levitt 1979: 26), which notably all show auxiliary variation in Sankoff & Thibault's 1977 paper on auxiliary alternation in Montréal French. In his *Grammaire de la langue française du seizième siècle*, Gougenheim (1951: 122) indicates that in the 16th century the choice between *être* and *avoir* with verbs of motion, such as *entrer* for example,[14] was not as restricted as in contemporary French: *avoir* was used to emphasise the motion itself and *être* to accentuate the completion of the motion. Marchello-Nizia *et al.* (2020: 836) include *entrer, périr* 'to die'*, broncher (tomber)* 'to fall'*, croistre* 'to increase, to grow'*, voler* 'to fly'*, couler* 'to flow, to sink'*, marcher* 'to walk'*, corre* 'to run, to hunt'*,* and *recorre* 'to rescue' in the list of motion verbs that could be conjugated with both auxiliaries in Middle French, using auxiliary *avoir* to indicate the action in itself and *estre* the action in its completion.

[13] Interestingly, this goes in the opposite direction of 2016 results, presented in §4.2.2.2. and §4.2.2.3., which show that inanimacy was one of the main significant factors influencing *avoir* choice.
[14] It is the only example given by Gougenheim (1951).

In fact, until the second half of the 19th century, most prescriptive French grammars allowed some motion verbs to be conjugated with both auxiliaries, depending upon whether the action was completed or not (Gougenheim 1951: 122; Sankoff & Thibault 1977: 82; Grevisse 2011: §782). In 1675, the grammarian Ménage (cited in Sankoff & Thibault 1977: 82) illustrates it in the following manner:

(4a) *Monsieur* ***a sorti*** (with *avoir*)
'Monsieur has gone out (but he has returned)'

(4b) *Monsieur* ***est sorti*** (with *être*)
'Monsieur has gone out (and has not come back yet)/is out'

However, in the spoken language, it remains quite unclear whether the French speakers of the day indeed made such aspectual distinctions. Marchello-Nizia *et al.* (2020: 1495) explain that it is probably not before the 17th century that the combination auxiliary *être* + past participle lost its resultative interpretation and that a semantic motivation for its restricted use with intransitive verbs disappeared. This is why more and more change of state verbs were able to combine with *avoir,* and the number of *être* verbs was therefore considerably reduced (Förster 1908: 69-103; Heidinger 2015: 282; Marchello-Nizia *et al.* 2020: 1495). According to Mackenzie (2006: 141-144), from the 18-19th centuries onward, *avoir* became the default perfect auxiliary while *être* stopped being productive.[15]

As to reflexive verbs, Grevisse indicates that they might appear with *avoir* in the *langue populaire* because Old French usages tended to "hesitate" between the auxiliaries *avoir* and *être* in the periphrastic tenses of reflexive verbs (Grevisse 2011: §782; Levitt 1979)[16]. However, Buridant (2000: §299) reports that in Old French and even in Middle French, in the

[15] Marchello-Nizia *et al.* (2020: 1495) note that today there are a few verbs with which the interpretation can be ambiguous, such as *il est mort depuis longtemps* 'he has been dead for a long time' (where *mort* appears to be a predicate participial adjective) and *il est mort il y a longtemps* 'he died a long time ago'.
[16] There is no mention of this "hesitation" at the Old French stage in Marchello-Nizia *et al.* (2020).

northeastern and Anglo-Norman areas, only a minority of pronominal verbs were conjugated with *avoir*. Stéfanini (1962) only reports around 3% of *avoir* use with pronominals in Old French.

Gougenheim (1951: 124) and Marchello-Nizia *et al.* (2020: 837) specify that in 16th century French, reflexive verbs were normally conjugated with *être* but reflexive constructions with a modal verb such as *il s'a voulu lever* 'he wanted to stand up' were common, and eventually gave way to *il a voulu se lever* in the 17th century (Gougenheim 1929: 184-185). However Gougenheim (1929: 183-184) points out that "*s'avoir pu faire* ['to have been able to do'] est barbare, le français n'admettant que l'auxiliaire *être* si le verbe est précédé d'un pronom réfléchi. *S'être pu faire* est donc la construction normale au Moyen-Âge, au XVIe et au XVIIe siècle."[17] As to the origins and reasons behind the use of *avoir* with pronominal verbs, Gougenheim (1929: 184) adds that "Les rares exemples de *s'avoir pu*, *s'avoir voulu*, viennent de l'usage dialectal et populaire d'*avoir* avec les verbes réfléchis, peut-être cependant aussi de l'usage habituel d'*avoir* avec les verbes auxiliaires de mode"[18], such as *vouloir* 'to want' or *pouvoir* 'may, can, to be able to'. According to Buridant (2000: §299), *avoir* with pronominal verbs is still often found in *français populaire*, but the examples to illustrate his claim are all excerpts from novels by French writer Henri Barbusse (1873-1935).[19]

While Grevisse (1969: 650) asserts that the conjugation of reflexive verbs with the auxiliary *être* (e.g. *Je me suis levé* 'I got up') was regularized under the influence of the construction copula + adjective (e.g. *Je suis levé* 'I am up'), Canale, Mougeon, & Bélanger (1978: 45) opine rather that for both reflexives and intransitive *É*-verbs, the selection of the

[17] '*s'avoir pu faire* is barbaric, since French only allows auxiliary *être* when the verb is preceded by a reflexive pronoun. *S'être pu faire* is therefore the normal form in the Middle Ages, in the 16th and 17th centuries.'
[18] 'The rare occurrences of *s'avoir pu*, *s'avoir voulu*, are found in dialectal and popular use of *avoir* with reflexive verbs, though maybe also from the conventional use of *avoir* with modal auxiliary verbs'
[19] Marchello-Nizia *et al.* (2020: 837) indicate that "in modern times" reflexive forms with *avoir* can now only be found [in fiction writings] in direct discourse to distinguish characters that are less educated.

auxiliary *être* in periphrastic tenses was simply standardized at the Modern French stage by prescriptive grammarians.

In order to explain the presence of French auxiliary alternation in the New World, it is worth taking into consideration the fact that Jones (2001: 109f., 2008, 2015: 43) has recorded various levels of generalization of *avoir* in both contemporary and traditional mainland as well as insular Norman French, such as in Jersey French (Jones 2001: 109f.), Guernsey French (Jones 2008), and Sark French (Jones 2015), especially with verbs *arriver, rester, (re)venir, mourir,* and *partir,* and with mainland Norman French being closer to standard French auxiliation than Insular Norman French (Jones 2015: 119, 121). Since there is no doubt that the Norman dialects carried over to Nouvelle-France by the settlers in the 17th-18th century strongly influenced Québécois French and that a few regional varieties of European French (Pooley 1988; Smith 2016)(see §2.4.1.), as well as many Oïl varieties (Smith 2016), such as Gallo (Deriano 2005), Picard (Auger 2003), and Walloon (Hendschel 2012) (see §2.2.1.2. for additional details and references regarding these varieties), display some extension of *avoir* in periphrastic tenses, it is legitimate to postulate that this morphosyntactic alternation in Laurentian French could have been inherited from these dialects, which at the time of the colonization of Nouvelle-France likely exhibited this variation (semantic or not) between the usage of *avoir* or *être.*[20]

2.1.2. From a prescriptive perspective

According to *Le Bon Usage* (Grevisse 2011: §782), the full list of intransitive verbs that are to be conjugated with *être,* in addition to the DR & MRS VANDERTRAMP verbs, traditionally includes *advenir* 'to happen'*, décéder* 'to die'*, échoir* 'to fall due' or 'to expire'*, intervenir* 'to

[20] For a more detailed diachronic account of auxiliary alternation in French, see Nyrop (1930: 212f.3), Levitt (1979), Tailleur (2007), and Rideout (2011, 2013).

intervene', *parvenir* 'to achieve', *provenir* 'to come from' or 'to be the result of', and *survenir* 'to take place', 'to arise', or 'to arrive unexpectedly'.

The European prescriptive conjugation grammar, *Le Bescherelle : L'art de conjuguer*,[21] indicates the note "*être* ou *avoir*" next to a few verbs, including *(r)entrer, (re)tomber, retourner* et *(res)sortir* to signal that it is a verb that is conjugated with both auxiliaries. However, the reader must refer to a separate chart, *Tableau 3* (Bescherelle 2006: 12-13), to understand when it is appropriate to use *avoir* or *être* with these verbs. The chart explains that certain verbs, not listed, are conjugated with *être* when they are intransitive (e.g. *Il est sorti de la salle à reculons* 'He backed out of the room') and with *avoir* when they are transitive (e.g. *Il a sorti son révolver de sa poche* 'He pulled out his gun from his pocket'). The same chart states that other verbs, again not listed, change their auxiliary depending on whether one insists on the action being done (e.g. *J'ai divorcé* 'I went through a divorce') or on the result of the action (e.g. *Je suis divorcé* 'I am divorced'). It is however difficult to determine with certainty whether the intentions of the speaker regularly match their auxiliary choice. The last section of the chart indicates that certain verbs belong to both categories: among the verbs studied in this research project there are *déménager*, *(re)descendre*, *(re)monter*, and *(re)passer*. One must therefore consult additional prescriptive grammars to determine which verbs must exclusively be conjugated with *être* when used intransitively, for example, *Le Bon Usage* (Grevisse 2011: §782). Among the intransitive verbs that only take *être,* the *Bescherelle* lists the same verbs as Grevisse, apart from *échoir*, but adds *obvenir* 'to revert to' and *ressortir (être du ressort de)* 'be under the jurisdiction of'.

Québécois prescriptive online tools such as the *Banque de dépannage linguistique* of the *Office québécois de la langue française* (OQLF) and the new dictionary *Usito* (Cajolet-Laganière, Martel, & Masson 2013), which aims to provide an open description of French for all francophones, are less reticent to acknowledge the variation. The OQLF indicates that

[21] Grevisse's *Le Bon Usage* and *Bescherelle : L'art de conjuguer* (informally referred to as *Le Grevisse* and *Le Bescherelle*, respectively) are both main reference works used in most Montréal francophone elementary and secondary schools.

déménager (like *commencer* 'to start', *descendre*, *embellir* 'to embellish', *empirer* 'to worsen', *rajeunir* 'to feel younger', *stationner* 'to park', *vieillir* 'to get older', etc.) can be conjugated with both *avoir* and *être* depending on the aspect: with *avoir* to emphasize the action of the verb, as in (5a), and with *être* to emphasize a state or the result of the action, as in (5b).[22]

(5a) *J'**ai déménagé** à l'automne dernier.*
'I moved (houses) last fall.'

(5b) *Je **suis déménagée** depuis six mois.*
'I moved (houses) six months ago/I have been living at my (new) place for six months.'

As for *passer,* the OQLF gives the same indication regarding the use of *avoir* or *être* depending on the aspect emphasized, but states that "en pratique cependant, *passer* s'emploie aujourd'hui le plus souvent avec *être*, du moins à la forme intransitive, c'est-à-dire sans complément d'objet. L'auxiliaire *avoir* n'est pas fautif et reste donc possible, mais il peut être senti comme vieilli."[23] With the verb *monter,* the OQLF is also quite descriptive and states that when the verb does not have a direct object, the two auxiliaries are in competition, and that in the past *monter* was conjugated with *avoir* to emphasize the action of the verb and with *être* to emphasize the result of the action, but that today tendencies differ. Interestingly, the OQLF explains that *être* is used more frequently than *avoir* when the subject of the verb *monter* is a human, as in (6a), but that *avoir* is also possible in certain contexts, and that when *monter* is used to refer to a level or a price, as in (6b), auxiliary *avoir* is more common.[24] The role of animacy will be discussed further in §2.4.2.3.11. and §4.2.2.2.

[22] http://bdl.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?T1=d%C3%A9m%C3%A9nager&btn_chercher=CHERCHER&id=3919 [webpage accessed on May 27, 2020]
[23] 'In practice however, *passer* is most often conjugated today with *être*, at least when used intransitively, namely without a direct object. The use of auxiliary *avoir* is not incorrect and is therefore possible, but can sound old-fashioned.' http://bdl.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?t1=1&id=4106 [webpage accessed on May 27, 2020]
[24] http://bdl.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?t1=1&id=3422 [webpage accessed on May 27, 2020]

(6a) *Simone n'**est** jamais **montée** dans un avion.* (or, less common: *Simone n'**a** jamais **monté** dans un avion.*)
'Simone has never been aboard an airplane.'

(6b) *Le thermomètre **a monté** à vingt-six degrés dans l'après-midi.* (or: *Le thermomètre **est monté** à vingt-six degrés.*)
'The thermometer went up to twenty-six degrees in the afternoon.'

The dictionary *Usito* simply states that intransitive (and transitive indirect) use of *passer* can be combined with *avoir* or *être,* but that *être* is more common,[25] and that intransitive *déménager* is used with both auxiliaries without a semantic difference.[26] As for *monter*, *Usito* claims that its intransitive (and transitive indirect) use can be with either *être* or *avoir*,[27] but that intransitive (and transitive indirect) use of *descendre* is with *être*, at times with *avoir* (but that the conjugation with *avoir* tends to go out of use).[28]

As for pronominal verbs, prescriptive works always dictate the use of auxiliary *être,* but Grevisse calls the selection of *être* with pronominal verbs a *servitude grammaticale* 'grammatical servitude' (Grevisse 1964: 5). He writes: "Pourquoi ne peut-on pas dire: « En tombant il s'a blessé » ? Pourtant il serait incontestablement logique de mettre là l'auxiliaire *avoir*."[29]

This brief survey of the historical and prescriptive literatures on French auxiliation has shown that these two perspectives are not particularly complementary nor enlightening enough to explain the reasons behind standard French auxiliation patterns and why auxiliary alternation in spoken French is being recorded all over North America (see §2.4.2. below).

[25] https://usito.usherbrooke.ca/d%C3%A9finitions/passer [webpage accessed on May 27, 2020]
[26] https://usito.usherbrooke.ca/d%C3%A9finitions/d%C3%A9m%C3%A9nager [webpage accessed on May 27, 2020]
[27] https://usito.usherbrooke.ca/d%C3%A9finitions/monter [webpage accessed on May 27, 2020]
[28] https://usito.usherbrooke.ca/d%C3%A9finitions/descendre [webpage accessed on May 27, 2020]
[29] 'Why couldn't we say: "He hurt himself when he fell' (with *avoir*)"? And yet, it would be unquestionably logical to use the auxiliary *avoir* in this case.'

The following section will show how syntactic and syntactic-semantic approaches have tried to shed light on the state of auxiliary selection in contemporary standard French and various other Romance languages, while the (variationist) sociolinguistic approach rather focusses on auxiliary alternation in spoken French and the fact that it is conditioned by several social and linguistic variables. The subsequent sections of this chapter appraise how the phenomena of auxiliary selection and alternation have been treated in these three fields.

2.2. Split auxiliary selection as a Romance phenomenon

Many Romance varieties show no alternation at all and generalize a single auxiliary, HAVE (reflexes of Latin *habere*), or more rarely BE (reflexes of Latin *esse*). Languages such as standard Castilian,[30] Portuguese, standard Catalan, Sicilian,[31] and many other Ibero-Romance and far southern Italian varieties (Ledgeway 2019: 349), as well as Aromanian,[32] now display extension of HAVE as the only auxiliary verb used to construct periphrastic tenses (Smith 1989: 311). By contrast, various central-southern dialects of Italy, such as Pescolanciano in Molise, have generalized BE instead (Tuttle 1986; Ledgeway 2019: 348).

Smith (1989: 311) suggests that a system requiring the use of two different auxiliary verbs is opaque, i.e. "is at variance with the norm for forms with a similar function", and this characteristic might favour the progressive extinction of BE as an auxiliary in Romance languages. It does not come as a surprise then that HAVE should be the sole auxiliary verb used to generate the "transparent forms" since it is the most unmarked and frequent auxiliary

[30] It has been suggested by Penny (2000: 48-51) that Castilian lost the alternation (the use of *ser* 'be') around the 15th century due to analogical levelling as well as dialect mixing during the Reconquista. Change of condition and location verbs like *morir* 'die' and *ir* 'go' were the last ones to give up BE in the 17th century (Legendre 2007: 169).
[31] In Sicilian and in Romanian, the forms of the lexical verb HAVE and the auxiliary HAVE are partly different: some forms of the auxiliary are phonologically reduced (shorter) forms of the lexical verb.
[32] The generalization of HAVE in Aromanian could be due to its considerable contact with Greek, which only exhibits auxiliary HAVE.

(Smith 1989: 311). The fact that two auxiliary verbs share one function is not transparent because from a functional perspective only one auxiliary is needed, as long as there no further functional distinction exists between them. And yet, this 'opaqueness' approach would suggest that split systems would not persist for millennia, even though they do in certain cases.

Amongst the Romance languages and dialects, there is a large amount of variation in terms of auxiliation. The most common distributional and structured patterns in *2-aux* 'two-auxiliary/split auxiliary selection' Romance systems are outlined here. Most of the studies conducted on Romance split auxiliary selection have highlighted how various languages and dialects display a grammatically structured distribution of the two auxiliaries, HAVE and BE. However, Ledgeway (2019: 348-349) mentions the seemingly unique case of the southern Calabrian dialect of Saline Ioniche, which displays free variation between HAVE and BE in all contexts (Manzini & Savoia 2005, II: 797-98, 800). Regarding free variation of HAVE/BE, Loporcaro (2016: 816) mentions that he knows "of no cases where it occurs across the board in the verb paradigm in all clause types", but adds that "the varieties coming closest to this are spoken in northern Apulia: Minervino Murge (Manzini & Savoia 2005, III: 27f.) displays full free variation in the pluperfect, whereas an auxiliation contrast persists in the present perfect just in the third singular, in contrast with free variation in all other persons exemplified with the first singular". It remains unclear whether social variables play a role in these alternations because studies of this kind have tended to convey little sense of the sociolinguistic dimension, as will be shown throughout §2.2.1. and §2.2.2.

Among the Romance languages that make structured use of two auxiliary verbs there are standard French and many Oïl dialects, Occitan, Catalan dialects, Sardinian, standard Italian and many Italo-Romance dialects, Raeto-Romance, and Romanian. Syntacticians have sought to explain how these languages structure their auxiliation.

2.2.1. Syntactic approaches to split auxiliary selection

Syntactic analyses of split auxiliary selection are based on evidence from split intransitivity, namely that the subject of BE-verbs (verbs selecting the BE auxiliary) shares properties with direct objects of transitives, an analysis supported by evidence from other constructions (impersonal passivization, verbs with reflexive clitics, etc.; see Aranovich 2007: 2), while the subject of HAVE-verbs (verbs selecting the HAVE auxiliary) shares them with the subject of transitives. This would entail that unaccusative clauses have an underlying object but no underlying subject (see §2.2.1.2.).

More evidence for a syntactic analysis of auxiliary selection comes from pronominal verbs in languages like Italian and French. In these languages, reflexive verbs (i.e. languages with reflexive clitics) select *essere*/*être* as the perfect auxiliary: the argument for a syntactic approach to split auxiliary selection (and split intransitivity) rests on the hypothesis that the subject of a reflexive verbs is an object at some abstract syntactic level. The existence of variation in split-auxiliary systems across languages forms a part of the evidence that syntactic theories of auxiliary selection use to argue against a semantic basis for the selection of BE or HAVE (Rosen 1984; Burzio 1986; Perlmutter 1989; Kayne 1993).

2.2.1.1. Syntactic mesoparameters: French as an outlier among Romance auxiliation systems?

Ledgeway (2019: 348) analyses auxiliation in Romance languages in terms of a syntactic predictive hierarchy involving four mesoparametric patterns of variation.[33] Among the

[33] Ledgeway (2019: 347) conceives linguistic variation in a scalar fashion modelled in terms of "parametric hierarchies":

> Macroparameters, the simplest and least marked options that uniformly apply to all functional heads, are placed at the very top of the hierarchy, but, as we move downwards, variation becomes progressively less 'macro' and, at the same time, more restricted with choices becoming progressively more limited to increasingly smaller subsets of features. […] More specifically, functional heads increasingly display a disparate behaviour in relation to particular feature values which may, for example, characterize: (i) a naturally definable class of functional heads (e.g. [+N], [+finite]), a case of mesoparametric variation; (ii) a small, lexically definable subclass of functional heads (e.g. pronominals, proper nouns, auxiliaries, unaccusatives), a case of microparametric variation proper; and (iii) one or more individual lexical items, a case of nanoparametric variation.

varieties that have a structured *2-aux* variation, auxiliary distribution appears to be sensitive to mood (§2.2.1.1.1.), tense (§2.2.1.1.2), person (§2.2.1.1.3.), or argument structure (§2.2.1.2.), as illustrated in Figure 2.1.

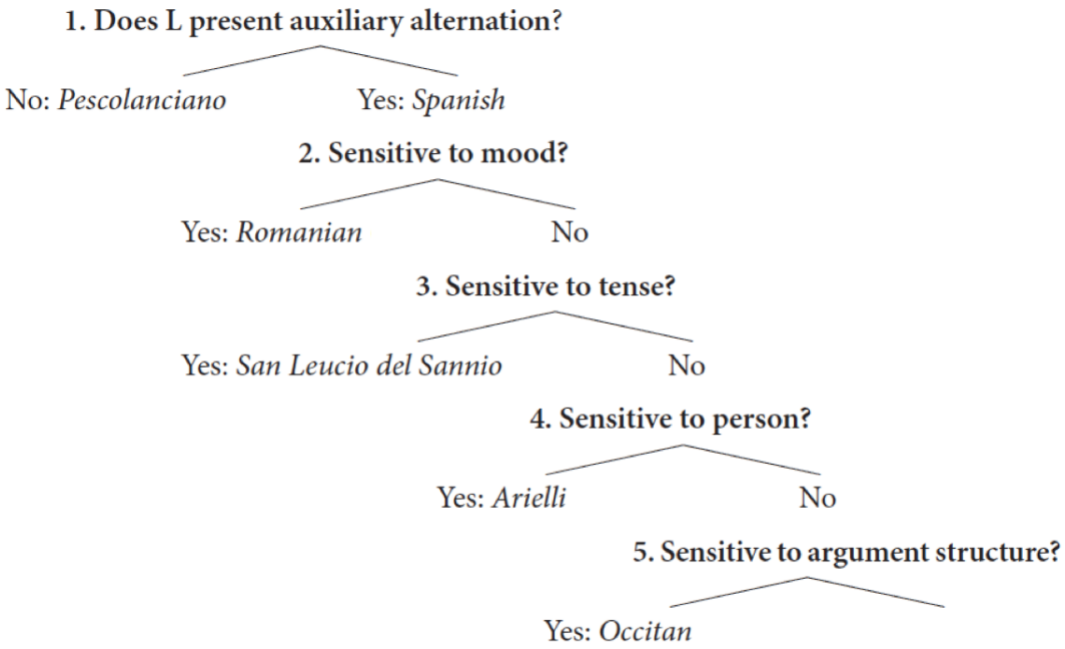

Figure 2.1 Predictive hierarchy of mesoparametric variation in Romance auxiliation (adapted from Ledgeway 2019: 348)

2.2.1.1.1. Mood

Romanian[34] stands as a perfect example of auxiliary distribution sensitive to mood, where auxiliary choice is entirely dictated by the realis (HAVE) vs irrealis (BE) mood distinction and encompasses distinctions of finiteness (with finite forms aligning with HAVE and non-finite forms with BE). The paradigmatic distribution occurs as follows: auxiliary BE (viz. *fi*) is found in the perfect infinitive (7a), the future and conditional perfect (7b), as well as the perfect subjunctive (7c), and the other auxiliary, always a form of HAVE (see footnote 31), is found

[34] See also Ledgeway (2014) for a detailed explanation of how auxiliaries BE and HAVE create two different syntactic structures in Romanian. However, the paradigmatic distributional difference is its most visible manifestation.

in the perfect indicative (7d) (Ledgeway 2019: 350). According to Smith (1989: 311-312), this distribution corresponds to a morphologically 'transparent' or logical/iconic system.

(7a) *Înainte de a fi mâncat / plecat citeam ziarul*
before of to be.INF eaten left I.read newspaper.DEF
'Before having eaten/left, I was reading the newspaper'

(7b) *Nu cred să fi mâncat / plecat*
not they.believe that be.SBJV eaten left
'They don't believe that I/you/(s)he/we/you/they have(/has) eaten/left'

(7c) *Vor / Ar fi mâncat / plecat*
they.will they.would be.INF eaten left
'They will/would have eaten/left'

(7d) *Am / Ai / A / Am / Aţi / Au mâncat / plecat*
I.have you.have (s)he.has we.have you.have they.have eaten left
'I/you/(s)he/we/you/they have(/has) eaten/left (…ate/left)'

(Examples taken from Ledgeway 2019: 350)

2.2.1.1.2. Tense

Sometimes auxiliary distribution is sensitive to tense. Ledgeway (2019: 351) analyses Sanleuciano (an Italian dialect of northern Campania, spoken in San Leucio del Sannio), where the present perfect and future-oriented conditional perfect/pluperfect subjunctive (counterfactual perfect) align with HAVE and the pluperfect indicative with BE. Another example of this pattern is the Salentino dialect of Torre Santa Susanna, where HAVE is used in the present perfect and BE in the pluperfect (Manzini and Savoia 2005,II: 794-95, cited in Ledgeway 2019: 350).

2.2.1.1.3. Subject person and number

Smith (1989: 311) discusses "alternative solutions" to split auxiliary selection that have been adopted by various Catalan and Italian dialects, which distribute the use of HAVE and BE according to the different grammatical persons (Tuttle 1986; Loporcaro 2007). It is worth

noting that the Romance varieties whose auxiliary distribution is sensitive to person are all non-standard and display this sensitivity solely in the present perfect tense (Ledgeway 2019: 357). Smith (1989: 311) also explains that one of the most common patterns involves conjugating the third person with HAVE, while the first and second persons take BE, and that these distributions are thus morphologically "transparent" (Lightfoot 1979: 121), rather than being syntactically or semantically "transparent" (Smith 1989: 311).

For instance, in Ariellese, an eastern Abruzzese dialect, the present BE is used for first and second persons and HAVE is employed for third persons (D'Alessandro & Roberts 2010, Ledgeway 2019: 351). This specific pattern of auxiliation is observed quite often because it distinguishes the basic discourse participants by their saliency: the first and second persons displaying a different form from the third ones (see Štichauer 2018 for a detailed analysis of person-based perfective auxiliation systems in Italo-Romance).

Ledgeway (2019: 364) points out that this pattern of person-based auxiliation in Italo-Romance could also be the result of the formal homophony of the 2SG and 3SG of HAVE in many dialects (viz. HABES/HABET 'have.2SG/3SG' > [a]), which could have led to the introduction of BE into the paradigm in order to disambiguate between persons. It should be pointed out that the disambiguation argument of homophony avoidance is hard to prove because very often homophony is simply not avoided.[35] As evidence, the Romanian BE subjunctive does not differentiate between persons (or number), in that the various persons are conjugated identically.

Loporcaro (2016: 813) notes that TAM-driven distributions of auxiliary selection "are prevalent across central-southern Italy, especially in mixed auxiliation systems" but that this parameter can be combined with person distributions:

[35] However, Andriani (2017: 230-231; personal communication) has recently suggested that homophony between 2SG and 3SG auxiliary plays a role in the paradigmatic distribution of progressive periphrases with *stare* and *andare* in Barese: the syncretism between 2/3SG present indicative 'st+a/v+a' ('stare/andare a' = STAND/GO + TO) has 'forced' the inflection to appear on the second conjunct of the periphrasis (the lexical verb) rather than on the auxiliary.

> For instance, in the Lazio dialects of Vallerotonda (Cocchi 1995:124) and Roiate (Orlandi 1989:66f.), BE has generalized to all persons in all tenses/moods, except the present perfect where HAVE remains in the third persons. This generalisation of BE may spare different persons: in Corese (province of Latina) HAVE remains only in the third person plural of the present perfect (Chiominto 1984:178-80), whereas in Zagarolo (province of Rome; cf. Lacetera 1982:112; Loporcaro 1999b:206-8) it continues in the third singular and plural of transitives and unergatives, but only in the third plural of unaccusatives (cf. Tuttle 1986:268). (Loporcaro 2016: 813)

According to Bentley and Eyþórsson (2001), from this point the pattern was then able to spread to other persons. Moreover, 1SG and 3PL frequently show formal convergence in auxiliary selection across Italian dialects: frequent formal homophony could once again be responsible for such a distribution, since 1SG and 3PL of the present BE become *so* in many dialects as well as in colloquial Italian. Another perspective on grammatical person is discussed by Croft (2003: 140-42), who observes that, globally, singular forms occur with a greater frequency than plural ones, and that third person forms are more frequent than first and second person ones. This could lead us to conclude that first and second persons would be rarer forms in natural speech, and one could expect them to undergo auxiliary generalization more easily than third persons. However, at the local level, grammatical person frequencies can also depend upon the different types of lexical verbs. Indeed, it has been observed that verbs of knowledge select first person pronouns more frequently (Smith 2011: 292-293).

Ledgeway also adds that with person-sensitive auxiliary selection, various layers of restriction (mood/tense, number, saliency) can be progressively superimposed or even combined. It is not clear whether any Romance auxiliation systems based simply on number exist, but the closest example might be the Pugliese dialect of Ruvo di Puglia, which displays 1SG. HAVE/BE, 2SG. B, 3SG. B, 1PL. H, 2PL. H, 3PL. H in the present perfect: this indicates that plural persons always align with HAVE and that BE is always marker of singular person.

Loporcaro's (2007) paper on triple auxiliation[36] systems in (Italo-)Romance focusses on language varieties that display a three-way contrast when it comes to perfective auxiliary selection in reflexive forms: rather than displaying a binary contrast (*2-aux* like in standard French) or showing no contrast at all (*1-aux* like in Spanish), these Romance varieties might display in reflexive forms free variation or variation sensitive to verb person.[37] Even though Loporcaro's work mostly makes use of Italo-Romance examples to explain his "triple" auxiliation systems, he claims that all the auxiliation patterns documented in Romance varieties can be mapped unto his implicational scale highlighting diachronic changes in the transition from *2-aux* to *1-aux* systems (for instance, in the history of Spanish (Aranovich 2003), Catalan (Batlle 2002), and Portuguese (Culbertson 2004)) (Loporcaro 2007: 210). Loporcaro concludes that "*3-aux* systems can be analysed as (diachronically unstable) intermediate steps towards the loss of the original *2-aux* rule" (2007: 211-212), which is what further investigation of auxiliary selection in pronominal forms in Montréal French could potentially show if auxiliary alternation in pronominals is observed, with or without a person-based pattern.

[36] Loporcaro's label of "triple auxiliation" is slightly misleading because it suggests that some (Italo-)Romance varieties have a third auxiliary verb, which is actually not the case.

[37] Loporcaro (2007) distinguishes between three different types of "triple" auxiliation systems, showing in reflexive forms free variation and/or variation sensitive to person: the Italian-Vallader compromise type, the Italian-Sardinian compromise type, and the Sardinian-Vallader compromise type. Frequently observed in northern Italy in the dialects of Veneto, Trentino, and Lombardy, the Italian-Vallader triple auxiliation system shows that all the reflexives pattern together and contrast both with unaccusatives, which take BE, and transitives/unergatives, which take HAVE. Moreover, all types of reflexives display "person-related alternation and/or free variation of auxiliaries E/H [BE/HAVE] in some or all persons of the paradigm." (Loporcaro 2007: 200). Loporcaro (2007: 202) explains that the Italian-Sardinian types, found in central and southern Italy, are more difficult to spot because they "are defined by contrasts observed in some persons only (even just in one), rather than in the whole paradigm". Interestingly, these systems also appear to be unstable because of this feature: "In dialects of this type, it is not rare for different groups of speakers to have different grammatical intuitions on the acceptability of auxiliaries, at least in some persons for some syntactic constructions." (Loporcaro 2007: 202). The last type of triple auxiliation system is found in southern Italy: in these varieties, free variation, which is not sensitive to person, affects only certain types of reflexives.

Not discussed in all these works, however, is whether the auxiliation of these languages behaves in the same way with subject pronouns and subject noun phrases, or whether this analysis only takes into account the grammatical person of the subject, regardless of its type.

2.2.1.2. Argument structure: the Unaccusative Hypothesis (Perlmutter 1978)

For Modern French, as for other Indo-European languages such as Italian, Dutch, and German, it has been argued that auxiliary selection correlates with unaccusativity (Perlmutter 1978; Burzio 1986). The Unaccusative Hypothesis specifies that there exist two types of intransitive verb: unaccusatives and unergatives, carrying different semantic and syntactic features (Perlmutter 1978: 160; Burzio 1986).

Semantically, an unaccusative verb is an intransitive verb whose grammatical subject is not a semantic agent, like the verbs *fall* or *float*. That is to say that it does not actively initiate, or is not actively responsible for, the action of the verb. In contrast, unergative verbs, such as *run* or *resign*, describe actions voluntarily initiated by the subject. Syntactically, with regard to auxiliary selection, it is assumed that in their periphrastic tenses, especially the perfect tense, the unaccusative verbs take the BE auxiliary, thus generally patterning with pronominals and passives,[38] and the unergative verbs require the HAVE auxiliary (Burzio 1986: 53; Loporcaro 2007: 187). The surface subject of unergative verbs behaves like the subject of a transitive verb, while the surface subject of unaccusative verbs acts rather like the object of a transitive verb (Burzio 1986: 30, 56, 74; Mackenzie 2006: 6). Thus, this theory suggests that while the underlying subjects of unergative verbs are also their surface subjects, the surface subjects of unaccusative verbs are in fact underlying objects (since this type of intransitive verb does not have underlying subjects). With unaccusative verbs, the subject is then not an agent of the

[38] Utheim (2013) shows that passive constructions correspond to an unaccusative structure. The same appears to be true for pronominal constructions (Buchard & Carlier 2008: 2426).

verb, in the sense that it is not responsible for the action of the verb, but rather undergoes it. Clear examples of unaccusative verbs in French include *tomber, mourir*, and *naître*. Mackenzie (2006: 117) notes, however, that the membership of certain French verbs in the unaccusative category does not necessarily determine their auxiliary, as opposed to Italian for example, so that obvious unaccusative verbs such as *exister* 'to exist' or *surgir* 'to arise' surface in the standard language with the auxiliary *avoir*. Auxiliary selection in French thus does not exclusively correspond synchronically to semantic or lexical features (Loporcaro 2007: 175).

Ledgeway (2019: 375) discusses auxiliary distribution sensitive to argument structure in (Italo-)Romance, which all Romance varieties once exhibited across all paradigms and persons. Various layers of restriction, such as mood (realis/irrealis), tense (past/present), number (singular/plural), and saliency (1/2SG), can also be progressively added or combined to this dimension (Ledgeway 2019: 364). A typical case of this auxiliation pattern displays a binary active-stative split: HAVE surfaces in conjunction with the subjects of transitive (A) and unergative verbs (S_A), and BE surfaces in conjunction with the subjects of unaccusative verbs (S_O). A case of this auxiliation pattern is Lengadocien Occitan. Ledgeway (2019: 365) explains that this pattern "represents the least constrained option and historically is also is [*sic*] the most widespread pattern in Romance which survives today in numerous Gallo-Romance and Italo-Romance varieties", where unaccusative verbs align with BE and transitive/unergative verbs align with HAVE for historical reasons (Vincent 1982; Ledgeway 2012: 130-34).

With regard to reflexive verbs, some auxiliation distinctions between subjects of non-reflexive unaccusatives and subjects of reflexive unaccusatives can be found (in Lower Engadinian, for instance). Ledgeway (2019: 370) adds, however, that this pattern can be more complex since reflexives can be divided into "at least four distinct classes ranging from the most unaccusative (inherent and direct transitive reflexives) to the more transitive (indirect unergative and indirect transitive reflexives) which may, in turn, align with differing auxiliation

alternations." For example, Loporcaro (2016: 817) observes that in northern Italy (e.g. central Veneto) non-reflexive S_O aligns with BE and A/S_A with HAVE, while reflexive S_O displays free variation.

Ledgeway (2019: 375) also discusses an additional dimension of variation with regard to BE-selection: unaccusativity being lexically fossilized, as in the case of many contemporary Oïl varieties. In such varieties, the selection of BE is now reduced to a "synchronically opaque, small number of intransitive predicates. This nanoparametric state of affairs, in which a once productive auxiliary distinction has all but fallen out of the system today precariously surviving in association with particular predicates as a lexical idiosyncrasy" (Ledgeway 2019: 375). Ledgeway cites the example of standard French, which as was noted above, retains now only between twenty and thirty unaccusative verbs that still select BE (Benveniste 1965: 181; Giancarli 2011: 373-74), representing only a small subset of the numerous verbs "which still systematically align with BE in varieties such as Italian, Occitan, or Corsican (Maiden & Robustelli 2007: 262; Giancarli 2011)", (Ledgeway 2019: 376). Additionally, it appears that 'popular' varieties of European French, and Occitan, such as Gascon (Bouzet 1963), display even fewer instances of BE-selection with most unaccusatives and in pronominals (Frei 1929: 86, 166; Bauche 1946: 105; Guiraud 1969: 40f.), as is the case in many eastern Oïl varieties (Remacle 1956: 39-48; Descusses 1986: 126; Hendschel 2012: 177 §166b, for Walloon). Moreover, some Picard dialects (Flutre 1955: 59; Vasseur 1996: 52; Dauby 1979: 35; Pooley 1988 for Picard French) also show retention of BE in just one or two unaccusatives, *mourir* and *aller*, and some generalize *avoir* even in pronominals (Auger 2003). Similarly, in the Lorrain variety of Ranrupt all unaccusatives take HAVE with the sole exception of *venir* which still licenses BE (Aub-Büscher 1962: 84 §107). Smith (1989: 320) notes that in some Occitan dialects, such as Limousin, Languedocien, and Provençal (as well as in Sardinian), if the subject follows the verbs, the auxiliary that surfaces is HAVE, and if the subject precedes the verb,

the selected auxiliary is BE (see Ronjat 1937; Roux 1895; Sicre 1909 for data on Occitan dialects; see Jones 1988 for Sardinian data).

Ledgeway (2019: 376) does mention the "free" alternation of HAVE~BE in unaccusatives and reflexives in certain Canadian French varieties,[39] but he notes that this variation can be partially attributed to "complex sociolinguistic factors (Sankoff & Thibault 1977; King & Nadasdi 2005; Rea 2014)". However, Ledgeway observes that in these varieties a subset of core unaccusatives (e.g. *aller*) tends to select BE more often than the others. Ledgeway (2019: 376) refers to all these cases as "synchronically unpredictable cases of lexical exceptions which residually reflect formerly more widespread and regular patterns of variation." It should be noted that auxiliary generalization in Canadian (and American) French varieties is also conditioned by several (socio)linguistic variables, which will be presented in §2.4.2.2., §2.4.2.3., §4.2., and §4.3.2. Using the work of Manente (2008: 42-43) as evidence, Ledgeway (2019: 376) claims that in the recent history of Québécois French[40] "HAVE has been extended to unaccusatives of change of location". Ledgeway's use of the verb "extend" is rather problematic here because it suggests that BE was once present or used to be found more commonly in Québécois French, for which there is yet no evidence. Moreover, Manente asserts that in Québécois French HAVE has "replaced" BE to mark punctual events, as in (8a), but that BE appears to have survived with these same verbs when they have a resultative reading,[41] as in (8b):

[39] This usage of the auxiliary *avoir* in contexts where, prescriptively, it 'should not' appear occurs as something notable in every variety of North American French, as well as in certain regions of France, and challenges the typology established by Zamboni (1998: 128, 2000: 86, 104f.), and later corroborated by Ledgeway (2012: 314), which divides Romance languages according to a northern-southern continuum. Among various structural features supposedly opposing northern language groups to southern ones, it serves us here to highlight the binary contrast in auxiliation between HAVE and BE, characteristic of northern languages (to which French belongs according to this typology) *versus* the generalization of a single auxiliary, either HAVE or BE, contingent "on the variety and/or the syntactic context" (Ledgeway 2012: 314).

[40] When citing the 'Québécois French' examples taken from Manente's research, Ledgeway (2019) sometimes refers to this variety simply as 'Canadian French', even though there exists another variety of Canadian French, Acadian French, which exhibits very distinct auxiliation patterns (see §2.4.2.).

[41] Interestingly, Romanian appears to behave similarly and has preserved what Ledgeway (2019: 376) calls a "relic" of auxiliary BE with a subset of unaccusatives, when these have a resultative value, as in (B) (Avram 1994: 494, 506-08; Motapanyane 2000: 16; Avram and Hill 2007: 49-52; Nevaci and Todi 2009: 142; Dragomirescu 2010: 210; Pană Dindelegan 2013: 228):

(8a) *Jean* ***a*** *arrivé/parti/entré/tombé* ***à huit heures/en deux minutes***
Jean has arrived left entered fallen at eight hours in two minutes
'Jean arrived/left/came in/fell at eight/in two minutes'

(8b) *Jean* ***est*** *arrivé/parti entré tombé* [***maintenant***]
Jean is arrived left entered fallen
'Jean is here/away/in(side)/on the floor [now]'

(Examples taken from Ledgeway 2019: 377)

However, Manente's claim about Québécois French are not particularly substantial because her Québécois French data appear to come exclusively from Sankoff & Thibault's 1977 paper, who has precisely shown with the 1971 data that speakers alternated, both individually and as a speech community, between *être* and *avoir* with sentences exemplified in (8a), and that auxiliary *être* was actually present in 66% of such tokens. Moreover, more recent work by Rea (2014) has shown that auxiliary BE is much more prominent with punctual event cases than in 1971. Since Manente also mentions the work of Canale, Mougeon, & Bélanger (1978) on auxiliary alternation in Ontarian French as additional evidence in this section of her thesis (2008: 42-43), it is also unclear whether she uses the term "Québécois French" as an equivalent for the umbrella term "Laurentian French", which would include at least both Québécois and Ontarian French (see footnote 3). This is also problematic since studies

---

(A) Ion **a / **e** sosit ieri / de ieri în oraş.
John has/is arrived yesterday/since yesterday in city
'John arrived yesterday/since yesterday in the city.'

(B) Ion **e /**a** sosit de ieri în oraş.
John is/has arrived since yesterday in city
'John has been here since yesterday in the city.'

(Examples taken from Dragomirescu and Nicolae 2009, cited in Ledgeway 2019: 376)

With these few unaccusatives in resultative readings, Dragomirescu & Nicolae (2009) believe the distribution of BE is not random and could have a semantic motivation: BE is selected by a subclass of unaccusatives situated at the top of Sorace's (2000) *Auxiliary Selection Hierarchy* (see §2.2.2.), i.e. "verbs of directed motion and change of location and verbs of (dis)appearance".

(Canale *et al.* 1978, among others) have shown that, in the 1970s, Ontario French had much higher rates of *avoir* selection in comparison to Québécois (Montréal) French.

2.2.2. Interface between syntax and semantics: The Auxiliary Selection Hierarchy (Sorace 2000)

Sorace (2000) introduces the Auxiliary Selection Hierarchy (ASH), an aspectual framework aiming to predict how Romance (mostly French and Italian) and Germanic languages (mostly Dutch and German) that have a two-auxiliary system distribute their perfective auxiliaries in intransitive verbs.[42] Verbs at the extremities of the hierarchy ('core' verbs, as opposed to 'peripheral' verbs in the centre of Table 2.1 which can display more variation) are change of location verbs, at the BE end, and non-motional process verbs, at the HAVE end.

Verbs at the extremities of the hierarchy (see examples 9 and 10) are characterized by the following properties: 'categorical and consistent syntactic behaviour across languages'; 'consistent behaviour within individual languages'; 'insensitivity to compositional properties of the predicate'; 'determinacy of native speakers' intuitions';[43] 'primacy in acquisition' and 'diachronic stability' (Sorace 2004: 256). The verbs that select BE tend to be unaccusative and the ones selecting HAVE, unergative:

[42] Her theory has also been further discussed and developed in Sorace (2004), Legendre & Sorace (2003), Legendre (2007), Legendre (2017), Ackema & Sorace (2017), among others.
[43] Loporcaro (2016: 818) notes that the study of Legendre & Sorace (2003) "propagates some dubious data such as It. *la popolarità del governo è scesa/ha sceso notevolmente* 'the government's popularity is/has dropped notably' and *sono/?ho rimasto solo* 'I am/?have remained alone'–judgments reported by Legendre & Sorace (2003:195f.), despite HAVE being totally ungrammatical in both examples in standard Italian."

The *Auxiliary Selection Hierarchy*	
Change of location	Selects BE (least variation)
Change of state	
Continuation of a pre-existing state	
Existence of state	
Uncontrolled processes	
Controlled processes (motional)	
Controlled processes (non-motional)	Selects HAVE (least variation)

Table 2.1 The *Auxiliary Selection Hierarchy* (Sorace 2000: 863)

Sorace (2004: 243) explains that the distinction is also systematically related to the semantic characteristics of the predicate: agentivity correlates with unergativity and patienthood correlates with unaccusativity (Perlmutter 1978; Dowty 1991). According to her (2004: 256), verbs such as *arriver*, *venir*, and *partir* (9a-c) belong to the top of the hierarchy and verbs like *travailler* 'to work' and *parler* 'to speak' (10a-c) belong to the bottom:

(9a) *Paolo* ***è venuto / *ha venuto*** *all'appuntamento* (Italian, I)
'Paolo is come/has come to the meeting'

(9b) *Ma sœur* ***est arrivée / *a arrivé*** *en retard* (French, F)
'My sister is/has arrived late'

(9c) *Maria* ***est / *at arrivata*** *a domo* (Sardinian, S)
'Maria is/has arrived at home'

(10a) *I delegati* ***hanno parlato / *sono parlati*** *tutto il giorno* (I)
'The delegates have talked/are talked whole the day'

(10b) *Les ouvriers* ***ont travaillé/ *sont travaillés*** *toute la nuit* (F)
'The workmen have worked/are worked whole the night'

(10c) *Los* [*sic: Sos*] *profesores* ***ont faeddadu / *son faeddados*** *totu su die* (S)
'The professors have talked/are talked whole the day'

(Examples taken from Sorace 2004: 256 and Legendre 2017: 279)

The 'core' verbs also select their auxiliary according to their inherent meaning, even when they appear alongside contradictory elements. This is illustrated in examples 11-13 (Sorace 2004: 257):

(11a) *Des plaintes* ***sont*** *arrivées continuellement* **atelic predicate**
'Complaints have arrived continuously'

(11b) *Des imitations* ***sont*** *apparues depuis des années*
'Copies have been appearing for years'

(12a) *Maria* ***est*** *tombée volontairement pour ne pas aller travailler* **agentive predicate**
'Maria has fallen on purpose in order not to have to go to work'

(12b) *Le verre* ***est*** *tombé de la table* **non-agentive predicate**
'The glass has fallen from the table'

(13) *Les policiers* ***ont*** *travaillé jusqu'à l'aube* **telic predicate**
'The police officers have worked until dawn'

Legendre & Sorace (2003) develop the hierarchy further in Table 2.2. based on the 'telicity' aspect and by showing the difference in the cut-off point where French and Italian select *avoir/avere*: for French it seems to be somewhere in the 'change of state' category, much higher in the hierarchy than it is in Italian. While in Italian, about a hundred verbs take BE,

French only selects *être* with twenty or so verbs that are a small subset of unaccusative verbs. (The auxiliary selection hierarchy is divided into seven classes and the * indicates variation.)

French	**Italian**	
BE	BE	**Change of location**: *arriver/arrivare, venir/venire,* etc.
		Change of state
BE	BE	a. Change of condition: *mourir/morire*, etc.
BE*	BE	b. Appearance: *apparaître/apparire* 'to appear', etc.
		c. Indefinite change in a particular direction:
BE*	BE	*monter/salire*, *descendre/scendere*
HAVE	BE*	*faner/appassire* 'to wilt', *empirer/peggiorare* 'to worsen', etc.[44]
HAVE	BE*	**Continuation of a pre-existing state**: *durer/durare* 'to last', etc.
		Existence of state:
HAVE	BE	a. *être/essere*
HAVE	BE*	b. *exister/esistere* 'to exist', *suffire à/bastare* 'to be enough'[45]
		Uncontrolled processes
HAVE	HAVE*	a. Emission: *résonner/risuonare* 'to resonate', etc.
HAVE	HAVE	b. Bodily functions: *suer/sudare* 'to sweat', etc.
HAVE	HAVE*	c. Involuntary actions: *trembler/tremare* 'to shiver', etc.
HAVE	HAVE*	**Motional controlled processes**: *nager/nuotare* 'to swim', etc.
HAVE	HAVE	**Non-motional controlled processes**: *travailler/lavorare* 'to work', etc.

Table 2.2 Auxiliary selection in (standard) French and (standard) Italian (adapted from Legendre & Sorace 2003: 200)

In terms of the historical applications of the ASH, Ackema & Sorace (2017: 12) explain that it is possible to "see its effects also in the diachronic development of those languages that have lost the split (or are in the process of losing it, as is possibly the case in varieties of French […])". According to them, the loss of BE starts with those verbs that are closest to

[44] It is my understanding that Sorace uses a metaphorical meaning of the word 'direction'.
[45] It is not clear to me whether these two examples actually signal the existence of a state.

the HAVE-taking verbs, at the bottom of the hierarchy, and the verbs at the top are the last affected by this loss (Ackema & Sorace 2017: 12). While this lexical-syntactic gradience process has been observed by Mateu (2009) in his work on Old Catalan and Old Spanish, which have both lost the split in their history, the only variety of French that is mentioned in this paper as being in the process of losing the split is Montréal French, with data taken exclusively from Sankoff & Thibault (1977). Ackema & Sorace's claim might therefore need to be slightly modified in light of the fact that Sankoff (2019) and this present study (see §4.2.1.) show that *être* is now much more prominent than in 1971.

I will not endeavour to form a detailed critique of the ASH because many scholars (including Bentley & Eythorsson 2003; Legendre 2007; Giancarli 2011, 2015; Utheim 2013; Kailuweit 2015; Loporcaro 2016; Legendre 2017) have already tried to improve and/or question the hierarchy. However, it is worth pointing out that one important flaw of the ASH is that it does not take into consideration pronominal verbs (Giancarli 2015: 86), even though they take BE both in Italian and in standard French (Legendre & Sorace 2003: 217). Moreover, the semantic-syntactic theoretical framework cannot fully explain French auxiliation, neither can it explain the immense variety of Romance auxiliary selection patterns based on mood, tense, and person, outlined in §2.2.1.1. The ASH cannot account for the use of *être* with verbs like *demeurer* and *rester* which are 'continuation of a pre-existing state' verbs. For Giancarli (2015: 83f.), telicity does not seem to play a role with regard to these two verbs, and Legendre & Sorace (2003: 213) noticed their paradoxical atelicity: "de manière inattendue, quelques verbes atéliques dénotant l'absence de changement sélectionnent *être*" ('unexpectedly, a few atelic verbs that denote the absence of change select BE'), but according to Giancarli (2015: 83), this should have made the scholars question "telicity as their main factor". He also points out that:

> telicity can only apply to verbs such as *monter* (go up) or *descendre* (go down) because it is understood in a very broad sense that denotes delimited events with an endpoint, such as *arriver* (arrive) (Sorace 2004: 246), as well as others expressing a direction without an endpoint, such as *monter* or *descendre* (Sorace 2000: 865). In Legendre and Sorace (2003: 212), these verbs are directly added to the list of exceptions. (Giancarli 2015: 83-84)

Indeed, the ASH does not explain why *monter* (14a) and *descendre* (14b) behave differently from other 'change of state' verbs which normally require the auxiliary *avoir*, like *exploser* 'to explode' (14c) and *fondre* 'to melt' (14d), for example:

(14a) *Je* ***suis montée*** *vers la colline.*
'I have gone up/climbed towards the hill'

(14b) *Je* ***suis descendue*** *vers la plaine.*
'I have gone down towards the plain'

(14c) *La bombe* ***a explosé****.*
'The bomb has exploded'

(14d) *La neige* ***a fondu****.*
'The snow has melted'

Furthermore, many scholars have recently been able to demonstrate through descriptive and quantitative variationist studies that in many areas of the *Francophonie* the *avoir* auxiliary is being generalized, but in varying lexical patterns. Before I summarize their main findings, the next section of this literature review adresses the theoretical debate that surrounded the nature of auxiliary alternation as a sociolinguistic variable.

2.3. The nature of auxiliary alternation as a sociolinguistic variable

During the 1970s, the decision by William Labov and Canadian researchers to extend the notion of sociolinguistic variable beyond phonology and phonetics gave rise to a major debate among linguistic scholars in France, the United States, and Canada (Sankoff 1972; Lavandera

1978; Labov 1978; Romaine 1981[46]; Godard 1992; Blanche-Benveniste 1997; Thibault 2001: 21; Deulofeu 2017).

This new direction was first criticized by Lavandera (1978), a sociolinguist and former doctoral student of Labov, and by the *Groupe aixois de recherches en syntaxe* (GARS) under the direction of Blanche-Benveniste who argued, against Labov (1978) and a group of Montréal sociolinguists (*L'école de Montréal*), that it was "inadequate at the current state of sociolinguistic research to extend to other levels of analysis of variation the notion of socio-linguistic variable originally developed on the basis of phonological data", because "the quantitative studies of variation which deal with morphological, syntactic, and lexical alternation suffer from the lack of an articulated theory of meaning" (Lavandera 1978: 171).

During these same years, the notion of syntactic variation in the study of spoken French was heavily discussed in meetings bringing together the Montréal sociolinguists and the linguists of the GARS. This led to a colloquium to discuss this specific topic, in May of 1979, where the main question ultimately underpinning the discussions was whether 'true' synonyms could exist in the study of a (morpho)syntactic variable (Blanche-Benveniste 1997: 19). Among the many syntactic variables that were quantitatively studied and analysed in these discussions, let us mention auxiliary alternation, but also negation exponence (*ne… pas* vs *pas*), the presence/absence of the subjunctive, the deletion of complementizer *que*[47], etc. (Blanche-Benveniste 1997: 19).

Taking the Sankoff & Thibault (1977) study on auxiliary alternation as an example, Lavandera suggested that its two authors should not have excluded the contexts where *être*

[46] For Romaine (1981: 17), the variable of auxiliary alternation did "not require a drastic alteration in the Labovian concept of linguistic variable because Sankoff and Thibault deal only with presence vs. absence of the variants in question." In her typology (1981: 15), a syntactic variable, such as the agentless passive, "suggests to me that a whole construction or arrangement of items which alternates is required" (Romaine 1981: 17), and shows no social or stylistic conditioning. And because Sankoff & Thibault (1977) had been able to divide their speakers into three groups: those who only use *avoir,* those who only use *être,* and variable users, Romaine (1981: 18) believes that "the case of *avoir/être* doesn't really constitute an example of syntactic variation as I would like to propose that it be defined because Sankoff and Thibault have completely ignored the problem of how the second and third groups express the aspectual distinction drawn by the first group, if indeed they do."

[47] See §3.2.3.16 for a list of studies on *que*-deletion in Québécois French.

and *avoir* are not interchangeable (such as when *être* behaves as a copula and the past participle of the verb as an adjective) because this denied them the possibility of highlighting the semantic differences between the two auxiliaries, which could precisely explain the variation (Lavandera 1978: 179; Thibault 2001: 22). For the Montréal sociolinguists, mostly influenced by an anthropological school of thought, there was no doubt that *J'ai resté à la maison hier* (lit. 'I HAVE stayed home yesterday') and *Je suis resté à la maison hier* (lit. 'I AM stayed home yesterday') meant the same thing, namely 'I (have) stayed home yesterday', and that the change in form could be explained by social differences (Sankoff & Thibault 1977; Blanche-Benveniste 1997: 19-20). Whereas the GARS team, influenced by Meillet and *l'école française de linguistique*, believed that to every difference in form corresponded a difference in meaning. In the case of auxiliary alternation, this implied that different social groups did not say exactly the same thing and, therefore, the way in which these groups understood each other was always approximate (Blanche-Benveniste 1997: 20).[48]

In a similar line of reasoning to Lavandera (1978), the GARS team decided to observe how the two auxiliaries had been treated in diachrony, in response to how Sankoff & Thibault (1977) had circumscribed the variable context (see §3.3.). They analysed what Littré and Vaugelas (citing Ménage) had written about the auxiliaries and how they had established semantic distinctions between *avoir* and *être* (Blanche-Benveniste 1997: 21). In these prescriptive works, it was said, for example, that *rester* and *sortir* should take auxiliary *avoir* when indicating that the action is completed and *être* when indicating that the action is still on-going (see §2.1.1.). For Blanche-Benveniste (1997: 21), the fact that Littré, Vaugelas, and Ménage specified an aspectual distinction dictating the distribution between *avoir* and *être* was sufficient

[48] However, Blanche-Benveniste (1997: 20) mentions that she has found similar *avoir* cases in France (Blanche-Benveniste 1977):
- *j'ai resté soixante-dix jours au lit* 'I (have) stayed seventy days in bed'
- *j'ai monté plusieurs fois à Paris* 'I have gone/went up to Paris many times'
- *oh ils ont pas sorti alors d'après ce que ton père m'a dit* 'oh they have not gone/did not go out from what your father told me'
- *et là j'ai pas resté deux jours chez eux j'ai resté une journée quoi pour leur faire plaisir* 'and then I have not stayed/did not stay two days at their house I (have) stayed one day you know to make them happy'

proof that it had been present in speech. However, these prescriptive comments could simply be cases of post facto rationalisation in an attempt to explain the variation between *avoir* and *être* that had been observed in speech at the time. As stated in §2.1.1., it therefore remains rather unclear whether speakers of the day actually made those aspectual distinctions. Moreover, Blanche-Benveniste cites Littré for the verb *tomber* ("*Tomber*, qui se construit d'ordinaire avec l'auxiliaire *être* dans les temps composés, peut aussi recevoir l'auxiliaire *avoir*."[49]), for which he allows both auxiliaries without any aspectual change (Blanche-Benveniste 1997: 21). She then suggests two hypotheses to explain the alternation observed in the 1970s, but indicates preferring the latter one: it would either be that "la différence de sens attestée à la fin du XIX[e] siècle s'est perdue, de sorte que l'alternance ne serait plus actuellement qu'une affaire de forme"[50] or that "tant que les deux formes ont été considérées comme correctes, on a accepté de légitimer une différence de sens. Depuis que l'une est devenue incorrecte, la tendance aurait été de lui dénier toute valeur sémantique."[51] (Blanche-Benveniste 1997: 21).

But for the Montréal sociolinguists, it was essential to exclude the contexts where *avoir* and *être* were not interchangeable in order to be able to discover any potential sociolinguistic influence on the alternation. With the variable data, it was noted that "these [aspectual] distinctions come into play neither in the intentions of the speaker nor in the interpretation of the interlocutor" (Sankoff 1988: 153). Moreover, the apparent nuances between synonyms seemed to Sankoff (1988: 154) "a posteriori artifacts of linguistic introspection", not unlike what happened to lexical items *deuxième*/*second* 'second/second and last', or orthographic

[49] '*Tomber,* which is normally constructed with auxiliary *être* in compound tenses, can also be conjugated with *avoir*.'
[50] 'the difference in meaning attested at the end of the 19th century was lost, so that the alternation would now only be a question of form'
[51] 'as long as the two forms were considered acceptable, we agreed to legitimise a difference in meaning. Since one of them has become incorrect, we would tend to deny it any semantic value.'

forms *philtre/filtre* 'philter/filter', which were subsequently attributed different meanings (Blanche-Benveniste 1997: 20).

To support the claim that different syntactic forms have different semantic interpretations, the GARS team and Blanche-Benveniste (1997: 25) use evidence from Damourette & Pichon (1911-1940) to claim that, for example, in "la pratique actuelle des Français" 'in the current speech habits of the French', negative marker *ne* sometimes served to orient the negation towards the verb rather than the adjective. The particle *ne* in the sentence *[Je ne la trouve pas] méchante* 'I do not find her nasty' would serve to differentiate it from sentence *Je la trouve [pas méchante]* 'I find her kind/okay/not bad'[52]. Sociolinguists, such as Sankoff & Vincent (1977), were however only interested in sentential negation and had therefore excluded tokens where it was the adjective that was negated. They would have been able to distinguish *[Je la trouve pas] méchante* from *Je la trouve [pas méchante]* because they have quite different prosodic contours.

The GARS also were of the opinion that different syntactic forms are not necessarily interchangeable because they can appear in various proportion depending on the context. Blanche-Benveniste (1977: 22) gives the example of an electrician: for a person who practices this profession, there is a difference between *Je vais au docteur* 'I go to the doctor (for a medical consultation)' and *Je vais chez le docteur* 'I go to the doctor ('s house to fix the electrical system)'. Blanche-Benveniste (1977: 22) therefore suggests that a variety of contexts must be studied, otherwise one could wrongly infer that the majority of speakers only use the form *Je vais au docteur*. However, what variationist sociolinguists aimed at studying was variation that took place within the same utterance context.

Blanche-Benveniste (1977: 23) made the same argument for variants associated with different levels of prestige, namely that they can appear in various proportion depending on

[52] More context is required to translate with more accuracy.

the speech style. Sociolinguistics did take this fact into account and, as an example, Sankoff & Vincent (1977) had discovered that the very few instances of negative marker *ne* in Montréal French appeared when speakers were discussing serious topics related morality, death, and religion, and did not claim that the variation observed in spoken French corpora was representative of all speech styles and contexts. Blanche-Benveniste (1997: 27) also recommends the use of 'very large' corpora of contemporary speech in order to discover whether two forms are actually interchangeable, but does not mention what she considers to be a 'very large' corpus. She gives the example of Calzolari (1995) who studied Italian verbs *chiedere* and *domandare* both meaning 'ask' for native speakers, but found that large corpora show only *chiedere* is used with complementizer *che*, even though both verbs supposedly allow it (Blanche-Benveniste 1997: 27).

The last argument used by the GARS team to invalidate the concept of syntactic variables as a sociolinguistic one is the presence of different geographical uses: if speakers from one speech community find variants from another speech community ungrammatical or "non françaises" 'not French' (Blanche-Benveniste 1977: 24), is it really right to talk about 'variation' in these cases? When sociolinguistic studies compare multiple speech communities, they take geographical differences into account in their analysis and specifically study how these impact speech patterns. Despite Montréal sociolinguists insisting on the fact that most of the syntactic variation they studied occurred within a single speech community, Blanche-Benveniste (1997: 28) argues that even though the GARS recognises that such variation can exist, it is rather with the interpretation of that variation that they disagree.

For Lavandera (1978: 171), who admitted like Blanche-Benveniste (1997) that conditioned variation exists at all levels of language, the question was rather where that specific tool, 'the sociolinguistic variable', stopped. Since Labov (1966) had presented evidence that:

> differences in form which had so far been analyzed as unmotivated and free, that is, referentially meaningless, were in fact carriers of some significance, social and stylistic, he provided specific evidence for the hypothesis that most if not all differences in form convey some information. (Lavandera 1978: 173).

As Lavandera (1978: 172) noted, in Labovian studies, when higher and lower scores of a variable are directly correlated with higher and lower positions on a socioeconomic scale or with higher and lower positions on a scale of formality of the context, those scores are interpreted as carriers of social or stylistic significance. Lavandera (1978: 173) therefore believed that it could now be demonstrated that differences in form could correlate with differences in meaning "once the notion of meaning was extended to include social and stylistic significance".[53] Lavandera's doubts about "the extension of probabilistic considerations from phonology to syntax" initially stemmed from Gillian Sankoff's 1972 paper "Above and beyond phonology in variable rules". In this paper, Sankoff had provided three examples of syntactic variation studied repeatedly by members of the Montréal group of sociolinguists, among which the deletion of the complementizer *que* in Montréal French. At the time of writing, Sankoff had reported no social or stylistic constraints on the data but observed "the presence or absence of *que* is differently allowable for different grammatical constructions" (Sankoff 1972: 54).[54] For Lavandera, this implied that the Montréal sociolinguists were studying a completely different type of variable because the study of variation, through phonological variables, had begun as the study of social and stylistic variation. She saw this as an issue because, according to Labov (1972b: 271), "social and stylistic variation presuppose the option of saying 'the same thing' in several different ways:

[53] In her opinion, quantitative studies reporting different scores for different age groups should therefore have been interpreted as carrying 'generational significance' or 'generational meaning', but instead generational differences had been interpreted as indicators of linguistic change (Lavandera 1978: 172).

[54] Further studies by Sankoff (1974, 1980a,b) have revealed a big social component in the variation of *que*-deletion, but to Lavandera (1978: 173) what mattered is that "for Sankoff's original proposal of undertaking the study of syntactic variation with the tools of phonological variation, she considered it sufficient to show evidence of grammatical complexity and of the influence of other syntactic processes on these alternating forms".

that is, the variants are identical in reference or truth value, but opposed in their social and/or stylistic significance". She clarified the difference between phonological variables and non-phonological ones as such:

> phonological variables which can be shown to have social and stylistic significance need not have referential meaning, while non-phonological variables are defined so that even when they do carry social and stylistic significance, they also have referential meaning, although this referential meaning is held to be the same for all variants. (Lavandera 1978: 176)

This sameness of meaning with syntactic variables is not obvious for Lavandera (1978: 175) as she explained that "units beyond phonology, let us say a morpheme, or a lexical item, or a syntactic construction, each have by definition a meaning. They are not like phonenes [*sic*] which, by definition, do not have any 'constancy of reference' as Sapir said." This would imply that variationists are never faced with exactly 'two ways of saying the same thing' when studying syntactic variables, and therefore Lavandera (1978: 175) questions whether that

> ground of clear semantic equivalence can be abandoned to carry out the same kind of study of variation for syntactic or morphological units which have to be proven to mean 'the same' to be treated as evidence of variability, and furthermore, whether semantic equivalence must in fact be a requirement at all.

For Labov (1978: 5), however, it was the case that for many syntactic variables there was "no problem whatsoever in establishing sameness of representational meaning. Thus negative concord, which has played such a large part in sociolinguistic work, is by definition multiple negation with the same truth value as single negation".

As it pertains to auxiliary alternation specifically, Lavandera (1978: 175) was of the opinion that this variable was a better sociolinguistic variable candidate than, say, the choice of active/passive in agentless sentences or the choice of *get/be* in the passive, because the referential meaning of *avoir* and *être* in these contexts is the same, and because Sankoff &

Thibault (1977) had shown that "in some contexts the choice of *avoir* instead of *être* is not promoted by the need to mean one thing instead of another, rather it is affected by the speaker's place in the linguistic market and by the different probabilities introduced by the lexical items of the main verb" (Lavandera 1978: 175). In order to analyse syntactic alternants as 'sociolinguistic variables', Lavandera (1978: 181) therefore insisted on the fact that two conditions be met:

> (1) that they can be proven to be the carriers of some non-referential information, to have social and stylistic or other significance, as is the case in the analysis of *être* and *avoir* but not in the analysis of the passive variable;[55] and (2) that they prove to be a kind of device of the language similar to phonological variables, that is, elements whose defining property is a quantifiable covariation and for which the frequency relationships are the very signals of those differences.

However, Lavandera (1978: 1979), in the same line of reasoning as the GARS team, believed that "unless we examine the entire distribution of the apparent synonyms, the possibility of explanation of the variation is ruled out. Sociolinguistic variables and variable rules, as they are defined, can only be heuristic devices, in no sense part of a theory of language." For Labov (1978: 2), this method was counter-productive: instead of extending meaning, he explains that one needs to "limit it much more narrowly than a formal linguist will do". Why?

> The answer is clear when we consider the simple demands of the everyday use of language. How do we know that someone talks like a countryman unless we know that there are rural forms and urban forms with the same meaning? How do we know that someone has spoken politely to us, unless we know that he chose one of several ways of saying the same thing, in this case the more mitigating variant. (Labov 1978: 2)

Labov (1978: 6) insisted on the fact that sociolinguists must not avoid the study of differences in meaning[56] and that circumscribing the variable context is precisely what sociolinguistic

[55] Labov (1978) and Godard (1992: 60) believed that a correlation with sociolinguistic factors was not a defining characteristic of variation.

[56] For Labov (1978: 8), lexical variables are slightly different from other non-phonological variables, even though they are treated in relatively the same way by Lavandera (1978). Labov (1978: 8) states:

analysis is about: "gradually isolating those cases where the same formal item has a different linguistic function and setting aside environments where the variation is neutralized or where the rule is categorized" can take a whole year of study. He added that "to return to the widest possible defining environment, as she [Lavandera] suggests, would reverse the process and lose the precision of analysis we aim at" (Labov 1978: 6).

But Lavandera (1978: 179) suggested analysing sociolinguistic variables with differences in referential meaning, even though she recognizes that "one of the reasons for restricting the study of variables to referentially meaningless surface variants is the fear of providing arguments which can be used irresponsibly to support ethnic, racial, and class-based prejudices". In other words, this type of evidence could be "used incorrectly to attribute to some groups the inability of thinking certain meanings" (Lavandera 1978: 180). For example, this would be the case if it were reported that, say, working-class speakers do not have the ability to use 'abstract' concepts as much as upper-class ones. For Lavandera (1978: 180), this would be the wrong conclusion to reach because "the evidence would not show an impossibility of using it, but a preference for the other forms. But more crucially, nobody has as yet proven that the kind of more general, or as Laberge [1977] calls it, more abstract meanings, reflect a cognitive or a communicative superiority".[57] Rather, the plausible conclusion, in her opinion, would be that "there are different conventional ways of

> In line with our general program of apportioning variance in linguistic choice among meaning and the various kinds of social significance, we will often encounter linguistic contrasts which potentially distinguish states of affairs but normally serve as social variants. Lavandera's own example of 'wiped out' vs. 'exhausted' reminds us that this is universally the case with lexical choice. We can prove that there are no true synonyms, in an absolute sense. But stylistic demands force us to substitute one word for another in speech and writing, so that in any given sequence of sentences we use many words as stylistic variants, though each has the potential ability to distinguish particular states of affairs.

[57] For Lavandera (1978: 180), the prejudice lies in "believing that since upper class speakers apparently make more use of these more referentially general linguistic categories (mislabeled 'abstract' [by Laberge (1977)]), these meanings can be taken as signs of a greater intelligence or of a more effective communication".

communicating 'the same' referential effect in the different sectors of the speech community" (Lavandera 1978: 179).

With regard to auxiliary alternation, how should one interpret the fact that for individuals who are situated at the top of the social scale auxiliary *avoir* is almost absent whereas it is much more prominent in the speech of lower-class speakers? Lavandera's argument would imply that people from a more privileged background would have at their disposal "une capacité de distinction aspectuelle dont les groupes sociaux moins favorisés n'ont pas l'usage"[58] (Thibault 2001: 22). Such a hypothesis would imply that the social distribution tends towards a complementary distribution based on social dialects but the verbs with which the upper classes use *être* systematically (such as *venir* and *aller*) are also those which yield the least *avoir* tokens in other social groups (Thibault 2001: 22; Sankoff & Thibault 1977). Conversely, there are verbs, such as *tomber* and *rester*, that are conjugated with *avoir* by all social strata, albeit at different rates (Sankoff & Thibault 1977; Thibault 2001: 22). For Sankoff & Thibault (1981), this complementarity was considered 'weak' and revealed a regular hierarchy of auxiliaries based on social classes and on verbs (Thibault 2001: 22). For variationist sociolinguists, these distributions simply denote the presence of a stable social stratification of the two auxiliary verbs, as opposed to cases of change in progress where different uses are observed according to various age groups (Thibault 2001: 22).

Godard (1992), who was a Research Director at *Laboratoire de linguistique formelle* of the *Centre National de la Recherche Scientifique* (France), also believed syntactic variation to be a bad candidate for variable rules.[59] Her main argument against the study of syntactic variation with variationist sociolinguistic methods is that the object being analysed by such rules is linguistic performance (the way a language system is used in communication) rather than linguistic competence (the system of linguistic rules/knowledge possessed by native speakers of a

[58] 'a capacity to establish aspectual distinctions that less privileged social groups do not have'
[59] Godard (1992) does not explicitly mention the study of Sankoff & Thibault (1977) in her paper. It is therefore possible that she does not consider auxiliary alternation to be a strictly syntactic variable.

language). This opposition was systematized by Chomsky (1965), who believed that only competence dealt with the linguistic system. According to Godard (1992: 55), quantitative studies of variation do not show that the linguistics system has variables rules, but that one can study *les règles de performance* 'performance rules' using quantitative and probabilistic tools. These ideas had been rejected by Labov (1972b: 226):

> One may set aside variable rules on the ground that they are rules of performance. The less said about this "wastebasket" use of the performance concept the better... The ability of human beings to accept, preserve, and interpret rules with variable constraints is clearly an important aspect of their linguistic competence or langue.

Labov instead aimed to incorporate regularities which were not usually taken into account in the grammar in the rules themselves, even though they are generated by a probabilistic model rather than a deterministic one (Labov 1972b: 226; Sankoff 1972; Cedergren & Sankoff 1974; Godard 1992: 56). Performance and competence are therefore blended into a single concept.

In a similar way to Blanche-Benveniste (1997), Godard (1992: 56) expresses concerns regarding *l'hypothèse de la compositionnalité* 'the principle of compositionality', namely that the meaning of a complex expression is determined by the meanings of each of its constituent expressions (lexical items) and of the (syntactic) rules used to combine them, such that two different structures carry different meanings. Moreover, Godard (1992: 57) noted some confusion in the definition of the syntactic variable in the debate opposing Lavandera (1978) and Labov (1978).[60] On the one hand, it can be two different syntactic sub-structures carrying the same meaning. On the other, they are variants that have the same 'truth' conditions, like Labov's passive variable. For Godard (1992: 59), the important question remains 'what does one want to study?':

[60] In the same article, Lavandera (1978) criticizes Weiner & Labov (1977) for treating active and passive sentences as variants because they do not carry the same meaning, but suggests that one should not limit the study of variation to forms that mean exactly the same thing (Godard 1992: 59).

> Si l'on cherche à mettre à jour les points d'hétérogénéité[61], les formes à comparer sont celles qui ont le même fonctionnement dans le système. Si l'on cherche seulement à préciser les conditions d'occurrence des formes en s'appuyant sur la comparaison de paires qui ont le même sens, alors il n'est pas indispensable que ces formes aient le même rôle dans le système.[62]

According to Godard (1992: 54), these two types of studies do not reveal the same thing about the linguistic system, and it would therefore not be a coincidence that Labov (1978), in his justification of the definition of alternation between active and passive sentences as a variable, "abandonne toute référence à la vision du système linguistique comme hétérogène pour insister sur la seule valeur heuristique de l'étude de la variation, qui permet une connaissance plus précise des propriétés des formes"[63]. For these reasons, Godard (1992: 64) is convinced that one must ultimately distinguish structural heterogeneity from the coexistence of different forms that carry the same meaning, and that the field of sociolinguistics is not suited for the study of syntactic variables.

Being followers of the Labovian approach and active members of *l'école de Montréal*, Gillian Sankoff and Pierrette Thibault, at the time when they were studying (morpho)syntactic variables, were undoubtedly very conscious of their role in this debate. It is therefore unsurprising that Sankoff & Thibault (1977) justify in so many details how they circumscribed the variable context of auxiliary alternation in Montréal French (see §3.3.).

[61] In a specific language, a system of rules that allows, in certain instances, the existence of competing subsystems with the same meaning. (Godard 1992: 60, 63)
[62] 'If one tries to uncover points of heterogeneity, the forms that ought to be compared are those that have the same function in the system. If one only tries to specify the conditions for occurrence of the forms by relying on the comparison of pairs with the same meaning, it is then not essential that these forms play the same role in the system.'
[63] 'relinquishes all references to the understanding of the linguistic system as heterogeneous, and rather insists on the sole heuristic value of the study of variation, which provides more accurate knowledge of properties of forms'

2.4. Recorded auxiliary alternation in French

2.4.1. European varieties

In addition to the various works of dialect description broaching the topic of auxiliary alternation in Oïl dialects that were mentioned earlier, other studies on European varieties of French have also shown that auxiliary alternation has been observed, among others, in the *français populaire* spoken in France (Bauche 1920; Frei 1929; Nyrop 1930: 212-213; Kukenheim 1967). In his work *Le Langage populaire. Grammaire, syntaxe et dictionnaire du français tel qu'on le parle dans le peuple avec tous les termes d'argot usuel* (1920), Bauche writes down sentences that he heard on the street, in the army, on the train, in shops, or that he had read in people's personal letters (1920: 27). He notes that auxiliary *avoir* often replaces *être* in "verbes neutres" 'neutral verbs' and pronominal verbs in the "langage populaire […] tel qu'on le parle dans le peuple de Paris"[64] (1920: 129). Bauche (1920: 129, 132) lists the following *avoir* examples with *monter, sortir*, *rentrer,* and *mourir* along with his 'translations' in standard French, in (15a-d).

(15a) *J'**ai monté** au deuxième* « je **suis monté** au second »
'I went up to the second floor'

(15b) *J'**ai sorti** le tantôt* « je **suis sorti** dans l'après-midi »
'I went out earlier/in the afternoon'

(15c) *Il **a rentré** ce matin* « il **est rentré** ce matin »
'He got back this morning'

(15d) *Il **a mouru** le 31 décembre* « il **est mort** le 31 décembre »
'He died on December 31st'

In *La Grammaire des fautes : introduction à la linguistique fonctionnelle, assimilation et différenciation, brièveté et invariabilité, expressivité*, Frei (1929: 86) discusses auxiliary alternation with *mourir* as quoted by Bauche in the previous example and specifies that this variant occurs

[64] 'The vernacular language as it is spoken by the people of Paris.'

"seulement dans le plus bas peuple, complètement inculte"[65]. Frei also notes a few examples of *avoir* use that he recorded in letters adressed to the *Agence des prisonniers de Guerre* (APG), an organization for war prisoners. Example 16 (from Frei 1929: 86) features auxiliary *avoir* with both *partir* and *arriver*.

(16) *Ils* ***aurait partie*** *sur le front le 18 novembre et ils* ***à arrivé*** *en Belgique le 27.* (APG)
'He/they would have left for the front on November 18 and he/they arrived in Belgium on the 27'.

Furthermore Bauche (1920: 129) records the use of pronominal *Je m'ai fait mal* 'I (have) hurt myself' while Frei (1929: 166) states that auxiliary alternation is similarly found in all subject persons in pronominal verbs. Kukenheim (1967: 65) also observes that the locution *Je m'ai lavé* 'I (have) washed myself' is attested in *français populaire*. More recently, Pooley (1988) recorded auxiliary alternation in the French dialect spoken in Roubaix, a city located in the Lille metropolitan area on the Belgian border in northern France.

2.4.2. North American varieties

As stated in §2.2.1.2, auxiliary alternation is far from being restricted to Montréal French and has been attested in other areas of the French-speaking world outside of Europe. Of relevance to this present study, in terms of methodology and comparative speech communities, are the variationist works listed below, which record and analyse auxiliary alternation in North American French in both its main dialects: Laurentian French, spoken mainly in Québec, in Ontario as well as in New England, and Acadian French, spoken mainly in the Maritime Provinces of Canada as well as in some remote regions of Québec (the Îles-de-la-Madeleine, the south coast of Gaspésie, and the Basse Côte-Nord), and in the American Madawaska.[66]

[65] 'only in the most vulgar people, completely uneducated.'
[66] http://continent.uottawa.ca/colloques-et-expositions/expositions/le-francais-au-canada-dun-ocean-a-lautre/le-francaisacadien/ [webpage accessed on June 24, 2020]

As Tagliamonte (2006: 5-6) puts it: "The essence of variationist sociolinguistics depends on three facts about language that are often ignored in the field of linguistics. First the notion of 'orderly heterogeneity' (Weinreich *et al.* 1968: 99-100), or what Labov (1982: 17) refers to as 'normal heterogeneity'; second, the fact that language changes perpetually; and third, that language conveys more than simply the meaning of words." As to the first point, sociolinguistic analyses of linguistic variation operate under the assumption that variability observed in a language is not random but actually structured. With this latter word, one should understand that speakers make decisions from among the different and available variable linguistic forms that are "systematically constrained by multiple linguistic and social factors that reflect underlying grammatical systems and that both reflect and partially constitute the social organization of the communities to which users of the language belong" (Bayley 2013: 85). The sociolinguistic variable in the Labovian sense is a set of alternative ways of saying the same thing, and the ultimate goal is to correlate these variants with independent variables: with social meaning (like social class, age, sex, ethnic group or contextual style) as well as with other aspects of language structure, and to derive rules for their distribution.

In Laurentian French, spoken by the majority of francophone Canadians, variation in auxiliary choice has been recorded in the province of Québec in Montréal (Sankoff & Thibault 1977, 1980; Thibault & Sankoff 1997, 2004 (unpublished); Sankoff, Thibault, & Wagner 2004 (unpublished); Rea 2014; Sankoff 2019), Québec City (Canale, Mougeon, & Bélanger 1977a), Ottawa-Hull (Willis 2000), Chicoutimi-Jonquière (Renaud & Villeneuve 2008), and in television interviews, with a range of influential or famous Québec personalities, where the participant's specific geographical origin is unspecified (Bigot 2011 for formal interviews and Villeneuve 2016 for informal ones). In the rest of Canada, the alternation has also been recorded in Ontario (Canale, Mougeon, & Bélanger 1978, in Welland, Sudbury Rayside-Balfour; Béniak & Mougeon 1989; Sammons, Nadasdi, & Mougeon 2015) and in the Western

Provinces (in Alberta by Papen 2016*[67] and Hallion Bres 2006*[68]; in Manitoba by Hallion 2000 and Hallion Bres 2006*). In New England, in the United States of America, Russo & Roberts (1999) recorded auxiliary alternation in Vermont French and Stelling (2011) recorded it in the French spoken in Massachusetts and in Rhode Island. In Acadian French, variation in auxiliary choice has been recorded in Prince Edward Island by King & Nadasdi (2005), in New Brunswick (Roussel 2016; Balcom 2008; Péronnet 1989), in Nova Scotia by Gesner (1978)*, and in Newfoundland by Brasseur (2000)*.[69] In Louisiana, in the United States of America, auxiliary alternation has been recorded in Cajun French, whose origin is Acadian French, by Papen & Rottet (1997)*, Picone & Valdman (2005)*, and Neumann-Holzschuh & Mitko (2018)*.

The Montréal French studies listed above have examined auxiliary alternation mostly in three main corpora from 1971, 1984, and 1995. This section will therefore describe how these relevant corpora were constructed and then provide a survey of all the (socio)linguistic variables that have been proven to influence auxiliary choice in the varieties of North American French enumerated above.

2.4.2.1. 50 years of Montréal French corpora

With regard to sociolinguistic research, Montréal is an exceptional case because in the last 50 years several corpora of spoken Montréal French have been created. For methodological (see §3.1. below) and comparative purposes (see §4.2.6. and §4.3.3. below), the corpora generally

[67] Starred works are not strictly speaking variationist studies but adopt a more descriptive approach.
[68] Hallion Bres (2006) has also observed auxiliary alternation in Mitchif French, a contact language.
[69] Auxiliary alternation has also been recorded in the nearby French *département français d'outre-mer* (DOM) Saint-Pierre-et-Miquelon archipelago (Brasseur & Chauveau 1990)*. Close to a thousand Acadian settlers started inhabiting the islands as early as 1763, following the *Grand Dérangement* of 1755 (the forced removal by the British of the Acadian people from Acadia, the present-day Canadian Maritime provinces of Nova Scotia, New Brunswick, Prince Edward Island, and northern Maine). Some Nova Scotia Acadians were deported to Virginia (USA) in 1755, then shipped to England in 1756, and to France in 1763, before being sent to Saint-Pierre-et-Miquelon. http://www.axl.cefan.ulaval.ca/amnord/stpierreetmiq.htm [webpage accessed on June 24, 2020]

referred to in this study are the Sankoff-Cedergren Montréal corpus of 1971, the Montréal 1984 corpus of Thibault & Vincent (1990), and the Montréal 1995 corpus of Vincent, Laforest, & Martel (1995), because all existing research on auxiliary alternation in Montréal French, with the exception of Rea (2014), has been conducted with data from these three corpora.

In 1971, Gillian Sankoff, David Sankoff, and Henrietta Cedergren co-designed and realized the original Montréal survey. As Sankoff & Cedergren (1972: 173-174) state, their goal was "a description of Montreal French in terms of basic grammar, most of which it shares with the standard language and other recent dialects of French, plus a unique configuration of features (phonological, syntactic, lexical, etc.) which characterizes Montreal French." As Cedergren explains (2018), they wanted to show that these distinctive features of Montréal French were elements of a "coherent system and not errors or abnormalities" shared by all members of the speech community, at a time of important linguistic alienation when the slogan *bien parler, c'est se respecter* 'to speak well is to respect oneself" appeared on posters all over the city, even on buses (Cedergren 2018; Blanc 1993: 246; Sankoff *et al.* 1976: 88-89). Sankoff, Sankoff, & Cedergren used a random stratified sampling method to obtain a sample of 120 speakers, by first locating Montréal neighbourhoods that were mainly French-speaking (by at least 64%) and by making sure that the participants selected from these areas were native French speakers who had resided in the city at least since the age when they began primary school (Sankoff 2019: 200). The three scholars had also studied the profile of different French-speaking socioeconomic areas, and civic addresses were selected using a random-numbers algorithm applied to a city address directory. The speaker sample was also stratified internally according to age (four groups of 30 speakers each) , sex (each age group contained 15 male and 15 female speakers), six socioeconomic (SES) levels, and geographical region. Each age/sex/SES category combination was hence represented by two or three speakers each. As Sankoff (2019: 200) further explains, "one fourth of the sample was

deliberately devoted to a small age range, from fifteen to nineteen years old, in the hope of adequately representing any new or incipient changes". The interviews, led by five native French-speaking university students, aged between 19 and 25 years old, lasted from half an hour to an hour and a half. As Sankoff (2019: 201) reports:

> The interviewers would knock on doors, starting with the address selected, until they located a willing subject who filled the requirements of the sampling grid. Explaining briefly that they were students, employed as interviewers on a study with the title of *Vie et coûtumes à Montréal* 'Life and customs in Montréal' […]. Interviews were not always one-on-one, because family members present in the home (usually spouses or parents, in the case of adolescents) were encouraged to participate if they wished. In terms of content, elicitation of demographic information was followed up by questions about childhood games, family celebrations of Christmas, and topics designed to elicit personal narratives (e.g. experiences of dating, injustices at school or at work) and views on fashion, religion, entertainment, and politics. A section on language asked interviewees to reflect on the status of French, and the interview ended with a short reading passage.[70]

In 1984, Pierrette Thibault and Diane Vincent (1990) carried out a follow-up study. They managed to locate and re-interview 60 participants from the original 1971 study and added a new group of twelve younger speakers aged between 15 and 25 years old (two male and two female speakers in each of three socioeconomic groups). The goal of this panel study was to allow for longitudinal analyses, and to find out more about linguistic change across the lifespan. Interviews varied in length, being mostly between about forty-five and one hundred minutes (Sankoff 2017: 25), and in degree to which respondents warmed to the occasion (Sankoff 2019: 201). Interviews in 1984 (and in 1995, see below) were also conducted by university students who were not previously known to the participants, and most of the conversation topics remained the same as in 1971. In 1984, respondents were also asked to recount what they had done over a recent day in their life (Sankoff 2019: 201).

In 1995, Diane Vincent, Marty Laforest, and Guylaine Martel (1995) undertook a

[70] Details of this procedure and additional information about the methodology of the study can be found in D. Sankoff *et al.* (1976) and G. Sankoff (2017).

second follow-up study. They located and re-interviewed twelve speakers of the previous 60 from the original 120 of the 1971 corpus, and two speakers of the twelve younger speakers who had been added in 1984. The three researchers became interested in discourse and conversation analysis, thinking that data obtained from regular variationist sociolinguistic interviews were not adequate for their research purposes. They recognized that the relationship between interviewer and interviewee was neither the most spontaneous nor frequent type of verbal interaction, and that the standard sociolinguistic interview was not well-adapted to studying the interactional component of informal discourse (Vincent *et al.* 1995: 29). They noted that a conversation between peers is normally characterized by quicker and more varied exchanges than the question/answer format of an interview. In order to specifically study this component of speech, they decided to resort to different ecological methods to obtain various speech styles and interactions, among which the use of self-recordings (Vincent *et al.* 1995: 29).

Table 2.3 schematizes the sequence of interviews across the three Montréal French corpora, including a total of 132 speakers (120 recruited in 1971, twelve recruited in 1984) and 206 interviews.

	1971		1984		1995	Total interviews
1971 speakers never reinterviewed	60					60
1971 speakers interviewed twice	48	→	48			96
1971 speakers interviewed three times	12	→	12	→	12	36
New 1984 speakers interviewed in 1995			2	→	2	4
New 1984 speakers not reinterviewed			10			10
Total speakers in each year	120		72		14	206

Table 2.3 Structure of the longitudinal Montréal French sample with data from 1971, 1984, and 1995. Arrows connect groups of reinterviewed speakers (adapted from Sankoff 2019: 200)

2.4.2.2. Sociolinguistic variables

Gender, age, and socioprofessional status (SPS) were taken to be the three basic macro-categories in first-wave variationist sociolinguistics, and their study is crucial to understand whether a variable is undergoing change (Sankoff 1988: 902). Many of the studies mentioned in the introduction to §2.4.2. have explored the impacts of sociolinguistic variables, but depending on the rates of *avoir* use, not all studies ascribe the same influence on auxiliary alternation to them. For example, Willis (2000: 104) concludes that sociolinguistic variables play a more important role in conditioning auxiliary choice than linguistic ones. By contrast Russo & Roberts (1999: 67) establish that external social variables do not play a significant role in the extensive replacement of *être* by *avoir* in Vermont French, which indicates "a later stage of linguistic change, supporting patterns found in the literature of language death", since it has been shown by Dorian (1978) "that variation may arise in endangered languages as a

result of the language death process, in the course of which simplified variants gradually replace more complex ones" (Russo & Roberts 1999: 70).

2.4.2.2.1. Gender of speaker

The potential effect of gender has been studied by all the variationist studies mentioned above, but was only found to play a role in Montréal French, and in Massachusetts and Rhode Island French. According to Sankoff & Thibault (1977) and Stelling (2011), male speakers use the non-standard form (conjugation with *avoir*) slightly more often than women. This evidence would be in line with the first half of the Gender Paradox, observed by Labov (2001: 293), in which "women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not."

2.4.2.2.2. Age of speaker

Younger speakers are known to be a vector of language change (Labov 2001: ch. 5), and Sankoff & Thibault (1977) as well as Canale *et al.* (1978) both observed that younger speakers tend to generalize *avoir* slightly more than their elders, which would potentially indicate a change in progress towards *avoir* at the time or an age-graded[71] stable variation. However, Sankoff (2019: 198) found that patterns of auxiliary alternation remained stable throughout an individual's lifespan so an age-grading interpretation could not be possible. Sankoff & Thibault (1977) believed that this change in progress away from the standard was constrained by pressure to conform to the standard. Willis (2000: 77) also observed more retention of *être* in older speakers in Ottawa-Hull and her results suggested a change in progress towards increased use of *avoir* with the *Ê*-verbs in Ottawa-Hull. This might be due to the fact that the

[71] An age-graded variation shows differences in speech habits within a community that are associated with age and occurs when individuals change their linguistic behavior throughout their lifetime, but the community as a whole does not change. For more details, see §2.5. and §3.5.1.

Anglicization of the Ottawa-Hull region is progressing to the detriment of French (Barrière 2009).

In contrast, Renaud & Villeneuve (2008) found in their study of Chicoutimi-Jonquière in the Saguenay region of Québec that older speakers were further away from the standard, namely they generalized *avoir* more often than young people. Stelling (2011) also found that New England French speakers[72] below the age of 70 were most likely to use *avoir*, and this suggested a change towards *avoir*.

2.4.2.2.3. Socioprofessional status/level of education of speaker

According to previous studies (Sankoff & Thibault 1977; Willis 2000; Stelling 2011; Sankoff 2019), speakers belonging to the lowest socioprofessional groups select *avoir* more often than the others, and these speakers are usually known to be a vector of language change (Labov 2001: ch. 5). In order to assign their participants to a class, Sankoff & Thibault (1977) used a six-level geographical scale based on the average income of the man of the household in the area of residence (Cedergren 2018), and later corpora (1984, 1995) made use of the following occupational scale from Thibault & Vincent (1990):

Upper and upper middle class
1. "Liberal professions" (e.g., law, medicine) and owners of businesses
2. Employed university graduates

Middle class and upper working class
3. Technicians, managers, foremen
4. White collar workers

Lower working class; lower class
5. Blue collar workers
6. No stable employment

[72] Stelling (2011: 1-2) defines "Franco-Americans" as the individuals whose families came from French Canada during 19th and early 20th century and settled in the northeastern USA. In 1900, Franco-Americans made up 60% of the population of the two communities under scrutiny (Southbridge, Massachusetts and Woonsocket, Rhode Island) and this proportion kept growing until the 1930s.

The first two levels of this scale combine occupation and level of education: most liberal professions are not fully independent of education (e.g. you cannot be a lawyer without qualifications) and the second level implies the completion of a university degree.

Willis (2000: 19) rather made use of two different scales, one for socioeconomic class (UNSK = unskilled workers and chronically unemployed, SK= skilled workers, SIS = sales and service, PROF = professional and managerial) and one for level of education (P = primary, S = secondary, PS = post-secondary), and demonstrated that less educated speakers tend to favour *avoir* more but did not mention whether socioeconomic class played a role in the alternation.

2.4.2.2.4. Linguistic Market Index

Developed by Sankoff & Laberge (1978) as an alternative way to look at socioprofessional status, the linguistic market index (LMI) adapts the notion of *marché linguistique*[73] from Bourdieu & Boltanski (1975), in which speakers are ranked according to a "scale that assesses the degree to which people have access to standard or legitimized language in their area of work" (Russo & Roberts 1999: 75). This index serves to palliate the fact that

> In many communities, people steer their speech toward more standard varieties than would be predicted of their socioeconomic status if their occupation warrants it. Indeed, this is so commonplace that some sociolinguists [Sankoff & Laberge 1978] have argued that the extent to which the standard language variety is valued in people's daily life – that is, their position in the linguistic market – plays a far greater role in shaping patterns of language variation and change than their class background. (Schilling 2013: 48)

Indeed Sankoff & Laberge (1978: 239) stipulated that their

[73] The concept of 'linguistic marketplace' was later expanded by Bourdieu in his book *Language and Symbolic Power* (1991).

> experience with the analysis of the [1971] Montreal French corpus [had led] to the realization that directly correlating linguistically variable behaviour with social class membership, whether defined statificationally or dialectically, [was] not a well-motivated procedure. It ignore[d] established facts such as that teachers, actors, and receptionists tend to speak a more standard variety than other people of similar social or economic position.

That is to say that the way in which a certain individual uses the language in their economic life tends to reflect their status in the broader social hierarchy (Bourdieu 1991). Sankoff & Thibault (1977) tested this by asking eight Québécois sociolinguists to rank the 120 speakers according to the importance of the legitimized language in their daily life. The LMI was found to be one of the best predictors to explain auxiliary alternation, since "speakers who figure importantly in the linguistic marketplace (i.e., whose work is somehow language-related: teachers, journalists, lawyers, and so on), whose linguistic production may indeed set standards as to what is considered "legitimate" – these people use *avoir* very little." (Sankoff & Thibault 1980: 340).

2.4.2.2.5. Language and dialect contact

In Ontario, Sammons, Nadasdi, & Mougeon (2015) showed that French-English bilingualism was actually a significant factor of influence for *avoir* selection with *aller* in the past infinitive (viz. *avoir allé*), and Béniak & Mougeon (1989) observed that language restriction played a significant role: highly restricted users of French showed a preference for *avoir*. Canale, Mougeon, & Bélanger (1977a: 3) suggested that "structural interference may take place in non-optimal areas of grammar where language internal simplification processes may already be at work" and that the use of the English auxiliary HAVE "may have served to reinforce or accelerate the leveling process that was already underway in Ontarian French" (Canale, Mougeon, & Bélanger 1977a: 3).

In contrast, Willis (2000) showed that fluency in English did not have an influence upon a speaker's auxiliary selection in Ottawa-Hull French. By classifying her speakers

according to the neighbourhoods in which they lived and in terms of reported proficiency, she calculated the role of English in terms of contact and found that her factor "Influence of English" played no role on the *avoir*-selection rates in Ottawa-Hull: speakers in neighbourhoods which had a closer contact with English did not produce more *avoir* tokens and the participants who displayed the highest and the lowest bilingualism ratings produced almost the same percentage of *avoir* in their *composé* utterances, with 59% and 62% *avoir* rates respectively (Willis 2000: 70-71).

With regard to contact with other French dialects, the results of Stelling (2011: 11) "lend support to the notion that contact with varieties of French other than the source dialect" discourages use of *avoir*: informants who were native French speakers but attended English school and learned French in the classroom only as a "foreign" standard language used *avoir* with much lower frequency (18%) than those who had French-English bilingual schooling (46%).

2.4.2.2.6. Stylistic variation

In his 2011 paper *Norme grammaticale du français parlé au Québec*, Bigot aimed to verify whether Québécois French had its own distinctive grammatical/syntactic standard, as opposed to a phonological one. His objective was to observe the social and cultural élite of Québec in a situation of formal communication in order to record their real grammatical usages. In order to build his corpus, he used televised journalistic interviews from a famous newsmagazine[74] on Radio-Canada (equivalent in status to the BBC in Britain) focusing on current affairs. He excluded interviews where the interviewee faced the camera, where more than one interviewee was present, and where the interviewee was not a native French-speaker originally from Québec. He analysed 14 variables in total, including the generalization of *avoir* with *Ê*-verbs.

[74] *Le Point,* which aired on televison from 1983 to 2006, explored the news in depth with interviews and documentary reports.

He found that only 5% (7/128) of his *Ê*-verb tokens surfaced with *avoir*, and concluded that standard oral Québec French thus resembles the *Bon usage* of Grevisse & Goosse (2008) very closely (Bigot 2011: 13). More recent work by Villeneuve (2016), using televised interviews discussing private matters[75], corroborates this result. While the type of data collected is slightly problematic because the context of a televised interview differs notably from a traditional Labovian interview, the fact that very little auxiliary alternation was reported does show that there was almost no deviation from the standard. However, because the number of tokens per speaker was so small, it is difficult to say with certainty that these results were representative of usage more broadly.

Using the frequency of use of French as a proxy for formality, Stelling (2011: 9) did however show that daily use of French was the only "frequency of use of French" variable that favoured use of *avoir* as an auxiliary verb in the speech communities of Southbridge, Massachusetts and Woonsocket, Rhodes Island. In these two speech communities, French is now used solely between family members and friends. For Stelling, it would seem that formality registers play a key role in conditioning the alternation. Conversely, Giancarli (2011: 14) included both spoken and written data from different registers (*soutenu*/*familier*/*populaire*/etc.) in his trilingual (Acadian French/Corsican/English) corpus, as well as theatrical plays and novels. However, he did not mention whether there was an effect of these various registers on the auxiliary data.

[75] *On prend toujours un train pour la vie*, which aired on Radio-Canada televison from 2008 to 2013, consists of conversations between host Josélito Michaud and his guests, aboard the Orford Express train, travelling between Magog and Sherbrooke. Guests, who were famous personalities and public figures in the early seasons of the show, usually share how they have dealt with or survived a life-altering event (death of a loved one, illness, accident, etc.).

2.4.2.3. Linguistic variables

2.4.2.3.1. Lexical effect

Sankoff & Thibault (1977: 96) noted that the various *Ê*-verbs did not select *avoir* in the same proportions, and created a ranking of verbs based on the selection rate of *avoir*. In order of probability of appearing with *avoir*, they are: *demeurer, rester, passer, déménager, tomber, rentrer, sortir, monter, descendre, retourner, partir, arriver, entrer, venir, revenir,* and *aller* (Sankoff & Thibault 1977: 96).[76] In the unpublished paper of 1997, Sankoff & Thibault concluded that in Montréal French, the distribution of *avoir* and *être* was both lexically and socially determined, partially semantically motivated, but was also to a considerable extent arbitrary.

Renaud & Villeneuve (2008) found a total of 24% of auxiliary alternation (excluding reflexive verbs) in the speech of 18 speakers in Chicoutimi-Jonquière, which varied among other things according to lexical verbs: *être* was always used with *décéder, mourir, naître,* and *aller*. *Être* was used more often than *avoir* with *entrer, (re)venir, devenir,* and *arriver,* and *avoir* was used the most with *passer, sortir, rentrer, (re)tomber,* and *monter*.

Canale, Mougeon, & Bélanger (1977a) compared the *avoir* rates of French-speaking students in Ontario with the auxiliary selection patterns of students from Québec City. They found that the ranking of the verbs which selected *avoir* in Québec City was roughly the same as that of French-speaking students in Ontario, but that the Québécois students displayed consistently lower selection rates of *avoir*. The percentages of *avoir* use are given first for the Québec students, followed by those for the Franco-Ontarians: *passer* (100%, 100%), *descendre* (67%, 91%), *tomber* (62%, 92%), *rentrer* (50%, 83%), *monter* (33%, 100%), *retourner* (33%, 75%), *rester* (used in the sense of 'remain', 33%, 82%), *sortir* (29%, 74%), *venir* (24%, 63%), *arriver*

[76] The verbs *naître* 'to be born', *décéder* 'to pass away', and *mourir* 'to die' also require the *être* auxiliary, but have been excluded from Sankoff & Thibault's 1977 study presumably because the past participles of *décéder* and *mourir* are most often used adjectivally. As to *naître*, it is possible that this verb has not shown any alternation in its auxiliary selection in the 1971 data. Alternation with *naître, mourir* and *décéder* has not been found in my corpus.

(10%, 60%), *partir* (9%, 41%), *revenir* (7%, 43%), reflexive or pronominal verbs (6%, 31%), *aller* (0%, 22%) (Canale, Mougeon, & Bélanger 1977a: 2).

Canale, Mougeon, & Bélanger (1978) analysed the *avoir*/*être* alternation in active periphrastic tenses in the spoken Ontarian French of Welland, Sudbury, and Rayside-Balfour. The researchers examined this variation within nine intransitive *É*-verbs in the spontaneous speech of 170 Franco-Ontarian students in French-speaking high schools of these three Ontario communities. They also noticed that even some reflexive verbs surfaced at times with the auxiliary *avoir* in Ontarian French. They observed that the verbs *tomber, rentrer, rester,* and *sortir* were conjugated exclusively with *avoir,* and the verbs *venir* and *arriver* showed strong tendencies in that direction (more than 50% of the utterances), while the verbs *partir, revenir, aller* as well as the reflexives verbs showed weaker tendencies of *avoir* selection (Canale *et al.* 1978: 48-49). In her study of Manitoba French, Hallion (2000: 365) recorded auxiliary alternation in verbs *arriver* (3%), *partir* (26%), *tomber* (59%), and *venir* (7%) with a lot of interspeaker variation, as well as in *passer, déménager, retourner, rester,* and *sortir*.

In Prince-Edward Island Acadian French, King & Nadasdi (2005) observed 99% *avoir* selection (*avoir* used almost always with *entrer, (re)venir, devenir,* and *arriver*), but some variation in *mourir* and *naître*. For *mourir*, it was observed that the auxiliary was correlated with the morphological structure of the past participle. If *avoir* was selected the past participle was *mouri*, but *être* appeared alongside the past participle *mort*. King & Nadasdi also noted that those two forms of *mourir* denoted two different aspectual interpretations: use of *avoir* was reserved for the act of dying and *être* for the state that resulted from it, comparably to what was described by Ménage (1675) in (4) (see §2.1.1.). Picone & Valdman (2005) echoed these conclusions and found that *avoir* may be used with all verbs, including intransitives and pronominals in Cajun French. Russo & Roberts' (1999: 83) argued that in Vermont the process of replacement might even be starting to affect the verbs that have traditionally shown

the strongest resistance to conjugation with *avoir*, such as the high-frequency verbs *aller, venir*, and *arriver*.

It should also be noted that every French variety studied exhibits a different number of verbs showing alternation, ranging from nine to twenty-one.

2.4.2.3.2. Verb meanings

Some of the intransitive verbs under study have multiple meanings, usually both a 'core' meaning as well as 'lexicalized'/figurative meanings. A few studies have also sought to determine whether their 'core' meaning (movement, change of state, etc.) selects *avoir* in the same proportions as their 'lexicalized' meaning. The examples of 'lexicalized' meaning in Sankoff, Thibault, & Wagner (2004) are *sortir avec quelqu'un* 'date someone' and *venir au monde* 'to be born' (lit. 'to come into the world'[77]). Thibault & Sankoff (1997) report that the effect of this factor was significant in variant selection with the verbs *partir, rester,* and *venir* (cited in Willis 2000: 39), where figurative use favoured *avoir* use. In Willis' study (2000: 62), the 'core' motional meaning of five intransitive motion verbs (*rentrer*, *tomber*, *partir*, *repartir,* and *venir*) and the 'core' continuing state interpretation of the stative verb *demeurer* showed the lowest percentage of *avoir*. Willis (2000: 60) also noted that when *sortir* was used to mean 'to date someone', it selected *avoir* 95% of the time, when used to mean 'to appear' (in the newspaper, etc.), it showed 88% of *avoir,* and when used to mean 'to go out (of a place)' it selected *avoir* with a rate of 74%. Since different meanings of the same lexical verb yielded different *avoir*-percentages, Willis (2000: 63) concluded that lexical semantics influences auxiliary selection.

In the combined 1971, 1984, and 1995 Montréal data, Sankoff (2019) analysed the differential rates of *avoir* in collocations involving *venir*, as presented in Table 2.4.

[77] In terms of compositionality, however, one could argue that *sortir avec quelqu'un* can have a 'core' reading because it can also mean "to get out of a place/go out in the company of someone". Additionally, *venir au monde* 'to come into the world/to be born' does not have to be read figuratively because "to be born" literally means to emerge into the world.

Collocation	N of *avoir* tokens	Total tokens *avoir* + *être*	% *avoir*
***Venir à bout* 'to succeed'**	13	31	42
***Venir au monde* 'to be born'**	15	119	13
***Venir (rouge)* 'to become (red)'**	3	30	10
***Venir* (movement only)**	13	438	3
Total	44	618	7,3

Table 2.4 Differential rates of *avoir* in collocations involving *venir* (adapted from Sankoff 2019: 205)

For Sankoff (2019: 205), *venir* stood as one of the most frequently used verbs, with a total of 618 tokens, and selected *avoir* infrequently, with an overall rate of 7.3%. This rate dropped to just 3% however when the three particular idiomatic expressions were removed, as shown in Table 2.4. The frequency of *avoir* was highest with *venir à bout* (lit. 'come to the end') at 42% (Sankoff 2019: 205). Similarly, in her study of Manitoba French, Hallion (2000: 365) observed that the verb *rester* when meaning *habiter* 'to live (somewhere)' was always used with *avoir*.

2.4.2.3.3. Frequency of use

In prior studies (Sankoff & Thibault 1977; Russo & Roberts 1999) it was shown that frequency of use impacts auxiliary alternation.[78] Since more frequent forms are less likely to undergo change (Martinet 1969; Bybee 2010), it is no surprise that recurrent verbs such as *aller*, *venir,* and *arriver* very rarely surfaced with *avoir* in Montréal French in 1971 (Sankoff & Thibault 1980: 334-335). Sankoff & Thibault determined the frequency of use of each verb based on their relative frequency in their corpus.

Kailuweit's 2015 paper on the residua of the semantic motivation behind the use of two auxiliaries in contemporary standard French analysed Sorace's ASH and concluded that other factors had to be taken into consideration in order to explain French auxiliation: change

[78] Frequency was also tested by Canale *et al.* (1978) and was not found to be statistically significant, but it is possibly due to the fact that they used a frequency dictionary of European French.

of location verbs which are at the top of the hierarchy, thus expected to take *être* categorically, also happen to be very frequent verbs (usually verbs of motion) and, according to Kailuweit, the parameter 'Frequency of use' of the various *Ê*-verbs would therefore be a better candidate than the ASH to explain, for example, why change of location verbs tend to select auxiliary *être* more than others (Kailuweit 2015: 272).

2.4.2.3.4. Possibility of parallel transitive use

Some of the verbs at issue also have transitive uses, and the existence of a parallel transitive use has also been considered to have a significant influence on auxiliary alternation because the transitive counterparts are homonymic (have the same spelling/pronunciation but different meanings) and always require the auxiliary *avoir*. According to Willis (2000: 35), "one can imagine that speakers may extrapolate from the fact that verbs with direct objects require *avoir* in the *passé composé*, and use *avoir* with those verbs capable of taking direct objects even in cases where they do not". Variationist studies mentioned earlier (Sankoff & Thibault 1977; Canale, Mougeon, & Bélanger 1978; Russo & Roberts 1999; Willis 2000; Renaud & Villeneuve 2008; Stelling 2011; Roussel 2016) have observed the correlation between the possibility of using some of these verbs transitively and the selection of *avoir* as an auxiliary when they are used intransitively. Amongst the verbs which have transitive equivalents, we can find: *(re)partir* 'to start up x (again)' (Québécism), *(re)monter* 'to bring (back) up x', 'to assemble x (again)', 'climb (back) up x', 'to pull up x', or 'to wind up x', *(re)descendre* 'to bring (back) down x', *entrer* 'make fit x' or 'key in x', *(res)sortir* 'to pull out x (from a place)' or 'to take out x (again), *(re)passer* 'to cross x (again)', 'to iron x', 'to spend (time)', 'to pass on x (again)', 'to (re)sit (an exam)', or 'to pass (an exam)' (Québécism), *rentrer* 'to bring in x', or 'to pull in x', *déménager* 'to move x',

and *retourner* 'to turn x over, inside out' or 'to return x'. The verbs which therefore do not allow parallel transitive use are *aller, arriver, rester, (re)tomber*[79]*,* and *(re)venir*.

Canale *et al.* (1978: 58) tested whether transitivity in reflexives could affect auxiliary selection in Ontarian French, as in *Je me suis coupé la main* 'I cut myself on the hand'[80] or *Je me suis acheté un livre* 'I bought myself a book', but found no evidence that transitivity played a role in the alternation. We also might expect to see more indirect reflexive constructions surface with *avoir*, based upon the fact that some of these reflexive verbs cannot be used intransitively. For instance, since one cannot say in French **J'ai lavé* 'I have washed' because *laver* is a transitive verb and requires an object,[81] is it the case that one is more likely to hear indirect constructions of the type *Je m'ai lavé les cheveux* 'I have washed my hair', rather than direct structures such as *Je m'ai lavé* 'I have washed myself'? Canale *et al.* also have analysed this alternation in terms of the existence of a lexical counterpart conjugated with *avoir*, since a pronominal *Ê*-verb such as *se lever* 'to stand up' may be said to have a lexical counterpart (though not a homophone) that is conjugated exclusively with *avoir* in the active voice *lever* 'to raise' (Canale *et al.* 1978: 58).

2.4.2.3.5. Possibility of parallel pronominal use

The existence of a parallel pronominal usage (*sortir* vs *se sortir* 'to get oneself out of, to escape') has been tested by Roussel (2016) as well as by Renaud & Villeneuve (2008), all of whom found that the effect of this variable was statistically significant in data from New Brunswick and Chicoutimi-Jonquière, respectively, in that it contributed positively to the probability of *avoir*. In her Ottawa-Hull data, Willis (2000: 65) however found the opposite effect, namely

[79] For future reference, it should be pointed out that the transitive use of *tomber*, meaning 'to beat', 'to seduce', 'to take off', or 'to drop' in popular Metropolitan French, is not attested in Québécois French (Sankoff & Thibault 1977: 99; Villers 2009: 1592).
[80] In some contexts, the sentence could also mean 'I cut my hand off'.
[81] A null object is possible in very specific cases, for example when referring to a previously mentioned item (Fónagy 1985; Larjavaara 2000): *J'ai lavé Ø ou j'ai pas lavé Ø* ? (my example).

that verbs which were not used reflexively in her corpus selected *avoir* with a higher rate. The list of verbs which appeared reflexively in her corpus include *s'arriver,*[82] *se passer, s'en redescendre, se retourner, s'en revenir, s'en sortir,* and *s'en venir* (Willis 2000: 38).

2.4.2.3.6. Possibility of parallel adjectival use

Another variable that seems to have some impact on auxiliary alternation is the availability of the past participle of the verbs for use as a resultative adjective. In order to test for the relevance of this variable, Sankoff & Thibault (1977) used a type of adjectival construction without copula, as exemplified in (17):

(17a) *Après minuit, les pensionnaires* ***sortis*** *ne peuvent plus rentrer.*
'Boarders out after midnight cannot get back in.' (Sankoff & Thibault 1980: 335)

(17b) ? *Un voyageur* ***arrivé*** *(d'un long voyage) est souvent content de retrouver son pays.*
'An arrived [*sic*] traveler is often happy to return to his own country.' or 'A traveler, arrived (from a long journey), is often happy to return to his own country.' (Sankoff & Thibault 1980: 334)

(17c) ***Un enfant* ***allé*** *(à l'école après une longue maladie) est souvent content de retrouver ses copains.*
'A child gone (to school after a long illness) is often happy to rediscover her friends.' (Sankoff & Thibault 1980: 334)

According to Sankoff & Thibault (1977: 99), the more acceptable the use of a past participle is as an adjective to express a state, the more it is possible to conjugate it with *avoir*. They add: "Les verbes qui résistent le plus à l'extension d'*avoir* dans leur conjugaisons aux temps composés sont précisément ceux qui n'admettent pas (ou n'admettent que

[82] An illustrative example of a *s'arriver* token, which to my ears is ungrammatical, reads as follows (Willis 2000: 117):
Puis je me suis arrivée dans un accident je me tu- tu-, ça me tuerait. (Speaker 067, line 980)
'Then I got into an accident I ki- ki- myself, - it would kill me.'

difficilement) les usages [- compléte]"[83] (Sankoff & Thibault, 1977: 99). Their results showed that the intransitive *Ê*-verbs that could have an adjectival use without copula were conjugated with *avoir* more often than the others. This is not entirely surprising given that Labelle (1992) included the possibility of parallel adjectival use as one of the various tests for unaccusativity, which is often correlated with auxiliary choice (see §2.2.1.2. above). For example, the verb *fondre* is unaccusative even though it takes the auxiliary *avoir,* as in (18):

(18a) *La neige* ***a/**est*** *fondu(e) pendant la nuit.*
'The snow has/is melted during the night'

(18b) *La neige* ***fondue***, *toutes les stations de ski ont fermé.*
'(With) the snow (having) melted, all the ski resorts closed down'

In some cases, the sentences in (17) were made acceptable[84] by adding what Sankoff & Thibault label "adverbial support", found in parentheses in (17b) and (17c) (e.g. *d'un long voyage* 'from a long journey' and *à l'école après une longue maladie* 'to school after a long illness'). Additionally, the influence of parallel adjectival use seemed to be greater than that of the existence of parallel transitive use in the findings of Sankoff & Thibault (1977: 336). Stelling (2011) hypothesized on the contrary that, by analogy, verbs whose past participles could be used as adjectives would be conjugated with *avoir* less frequently because adjectives are generally not preceded by *avoir.* In his data, the effect of this variable did not prove statistically significant. For Canale *et al.* (1978: 51), the possibility of using the past participle adjectivally was tested with copula (exemplified in 19) and was only acceptable with *sortir, partir, arriver, rentrer, revenir*, and *tomber,*[85] and not with *aller, rester, venir,* and the reflexives:

[83] 'The verbs that resist the most to *avoir* generalization in the conjugation of their periphrastic tenses are the ones that precisely do not allow (or allow with difficulty) the [- completed] uses.'
[84] Sankoff & Thibault (1977) do not mention which Montréalers participated in this grammatical judgement task, determining which verbs could be used as adjectives, and how the task was put together.
[85] As a native speaker, I am not convinced that the past participle of *tomber* can have a stative reading.

(19) *Marie* ***est sortie/partie/arrivée/rentrée/revenue/tombée*** *maintenant.*
'Mary is out/gone/here/back/?fallen now'

The ability or inability to be used adjectivally with the copula did seem to correlate with the percentage of *avoir* or *être* use in the cases of *aller, arriver, rentrer, sortir, tomber,* and the reflexives (Canale *et al.* 1978: 52). However, this variable alone could not account for the high frequency of *avoir* with *rester* and *venir,* nor the high frequency of *être* with *partir* and *revenir* (Canale *et al.* 1978: 52).

For Willis (2000), as well as for Roussel (2016), admissibility of parallel adjectival use (combined with transitive use for Roussel 2006) was the greatest determinant of variant choice. To code this variable, Willis (2000: 36-37) looked at actual speaker usage and examined each of the *Ê*-verbs in her Ottawa-Hull corpus to determine which past participles had also been used as adjectives (most were used with copula). The ones which had appeared as adjectives in her corpus were *commencer, décéder, déménager, demeurer, monter, partir, passer, rentrer, repartir, rester, retourner, revenir, sortir,* as well as *tomber*, and those which did not were *apparaître, arriver, descendre, devenir, entrer, redescendre,* and *venir.*

2.4.2.3.7. Morphologically-derived forms

According to Kiparsky (1973), morphologically-derived forms often seem to be regularized more easily than non-derived ones. In their study of auxiliary alternation in Ontario, Canale *et al.* (1978: 53-54) cited the Canadian English example of the non-derived form *cast* which continues to surface as *cast* in the past tense, but its morphologically-derived forms *broadcast* and *forecast* tend to show a regularized past tense in *-ed*, i.e. *broadcasted, forecasted.*[86] Canale *et al.*

[86] This is not the case in standard British English, where *broadcast* and *forecast* are the normal past forms on, say, the BBC. It is also worth noting that they are high-frequency items and so potentially less likely to undergo analogical change. This claim is based on a simple Google search of ["rain is forecast" site:www.bbc.co.uk] yielding 639 results vs ["rain is forecasted" site:www.bbc.co.uk] yielding 2 results and of ["show is broadcast" site:www.bbc.co.uk] yielding 574 results vs ["show is broadcasted" site:www.bbc.co.uk] yielding 0 results, as of April 11th 2017.

(1978: 54) tested whether morphologically-derived forms (e.g. *revenir*) behaved in the same way as non-derived forms (e.g. *venir*) and, on the contrary, they found that morphologically complex verbs showed resistance to the spread of *avoir*: for instance, *revenir* was conjugated much less frequently with *avoir* than *venir* in their data, suggesting rather that basic and derived forms are simply liable to behave differently in morphological change. Again, this is most probably because change occurs first in unmarked contexts (Smith 1999; Andersen 1990; Stein 1989; Timberlake 1977). For Canale *et al.* (1978: 54), pronominal verbs are themselves morphologically-derived (e.g. *sortir* > *se sortir*, my example) and, in their study, they did not appear to be as easily affected by this levelling process as other non-derived verbs.

2.4.2.3.8. Subject person and number

Renaud & Villeneuve (2008) found that auxiliary alternation varied according to subject pronouns: *avoir* was most often selected with the 3PL person and with the 1SG person *je*. For Sankoff (2019: 207), a subject in the 3SG person *il/elle* 'he/she' ($p < 0.001$) and *on* (usually referencing plural 'we') ($p < 0.01$) statistically disfavoured use of *avoir*. With pronominal verbs, Hallion (2000: 357) observed that *avoir* selection was most common with the 1SG person (15/21 tokens, 71%). With restricted speakers of French, she noted that with *venir* tokens the 3PL person of the *passé composé* was always conjugated with *avoir* (12 tokens out of 12) (Hallion 2000: 367).

2.4.2.3.9. Tense and mood of the verb

Sankoff (2019: 207) tested the influence of tense and her results were statistically significant: when the token occurred in tenses other than the *passé composé* it favoured use of *avoir*. She does not mention however whether she looked at the separate tenses first or merged them

and analysed them together against the *passé composé*. In contrast, Roussel's (2016) data from the *Corpus de français acadien du Nord-Est du Nouveau-Brunswick* (Beaulieu 1995) showed that the present tense of the auxiliary (*passé composé*) affected variant choice in favour of *avoir*. Renaud & Villeneuve (2008) found that verb tense influenced auxiliary alternation: more *avoir* use with *passé composé* than with *plus-que-parfait*.

2.4.2.3.10. Type of subject

Renaud & Villeneuve (2008) observed that auxiliary alternation varied according to the types of subject: *avoir* was most often selected with indefinite *ce/ça* 'this/that'. Sankoff (2019) also found *ce/ça* to statistically favour the use of *avoir,* but also noted that *qui* 'who/whom' as the subject of a relative clause disfavoured *avoir* use.

2.4.2.3.11. (In)animacy of the subject

Thibault & Sankoff tested this variable in their 1997 (unpublished) paper by looking at the 1971 and 1984 Montréal data, but the results however were not statistically significant, except in the case of *partir* 'to leave', where inanimate subjects favoured *avoir*. Sankoff (2019) re-tested this data and then found the effect of this variable to be statistically significant: *avoir* was more likely to be selected when the subject was inanimate than for human or dummy subjects ($p < 0.01$). This is consistent with the principle that linguistic change tends to occur first in unmarked contexts (Smith 1999; Stein 1989; Timberlake 1977): if the diachronic trajectory is now towards *avoir*, it might be the case that inanimate and impersonal subjects are potentially more likely to favour *avoir* since they are less marked (i.e. they are the default option, are more common, basic, regular, etc.). This variable was also tested by Willis (2000: 66) but was not found to have a significant effect.

2.4.2.3.12. Presence of intervening element between the auxiliary and the past participle

Elements can intervene between the auxiliary and the past participle and are usually adverbs (*Il* ***est*** *tombé* vs *Il* ***a*** *vraiment tombé* 'he has fallen hard'), including adverbs of negation (*Il* ***a*** *pas tombé* 'He did not fall'). The intervening material can also be comprised of more than one element, like a sequence of adverbs. This variable was tested by Thibault & Sankoff in an (unpublished) 1997 paper but the results were only significant for the verb *rester*. Sankoff (2019: 207) re-tested this data and found this variable to have a statistically significant effect: *avoir* was more likely when there was intervening material (usually adverbial[87]) between auxiliary and past participle ($p < 0.01$). The adjacency of past participle and auxiliary was tested by Willis (2000: 65), who reported that non-adjacency favoured *avoir*.

2.5. Real time and apparent time

Before the development of sociolinguistics as a field of study, linguistic change was traditionally observed in 'real time' by comparing two (or more) points of the history of a language on a timeline: structural linguists believed that it was the only way that "changes could be observed because apprehending them while they were in progress was theoretically impossible" (Chambers 2013: 307) (see Bloomfield 1933: 347; Hockett 1958: 444; discussed critically in Labov 1972: 21-23, 1994: 44-45; Chambers 2009: 198-200). However, Labov (1963, 1966) developed[88] methodological techniques in his works on Martha's Vineyard and New York City that allowed linguists to track linguistic changes while they were happening and consequently "established the basis for a synchronic approach to language change", among which the apparent-time construct, "a surrogate for the real-time examination of data"

[87] Sankoff (2019) does not mention whether the intervening material can be anything other than adverbs.
[88] Labov (1966: 278, 391) points out that he was not the first one to use differences between generations to make inferences about diachronic change. In 1905, Gauchat had used apparent time to study sound change in the Swiss village of Charmey, and Hermann's (1929) reinvestigation of the same village approximately 30 years later essentially confirmed Gauchat's results (Bailey *et al.* 1991: 241; Sankoff 2006: 110; Cukor-Avila & Bailey 2013: 259).

(Cukor-Avila & Bailey 2013: 239, 240). The study of change in progress, through the observation of intergenerational differences at a given point in time, is based on this apparent-time construct. It hypothesizes that linguistic differences between generations (apparent-time differences), with stylistic and social factors being held constant, mirror diachronic changes in the language (real-time linguistic developments). This is thought to be the case because, for the most part and under normal circumstances, "the core features of an individual's vernacular language variety are solidified for life by the time they reach their late teens"[89] (Schilling-Estes 2005: 219). Crucially, Labov's comparison of apparent-time distributions with real-time evidence, collected some thirty years earlier for the Linguistic Atlas of New England (Kurath, Bloch, & Hansen 1939), corroborated his arguments: the increased use of centralized onsets of diphthongs (ay) and (aw) on Martha's Vineyard in younger generations compared in apparent time to older ones, reflected an observed real-time diachronic increase in the use of these features (Labov 1963).

The apparent-time construct relies on the assumption that, in the vast majority of cases, individual vernaculars remain fairly stable during the lifetime of an adult; however recent data suggest that they can behave in somewhat more complex ways (cf. Sankoff & Blondeau 2007; Cukor-Avila & Bailey 2011; Wagner & Sankoff 2011; Sankoff & Thibault 2011; Cukor-Avila 2012; Sankoff 2004, 2019). With apparent-time studies, there are two situations where data showing differences across age groups cannot be assumed to represent actual diachronic linguistic changes, and confirmation in real time is therefore necessary. These include changes throughout the lifespan of an individual and cyclical age-grading (namely when differences in speech habits are associated with specific age categories or particular life phases but where the speech community as a whole does not change) (Labov 1994: 84). Moreover, very specific methodological steps have to be followed for real-time

[89] According to Chambers (2013: 318), the formative years for dialect and accent formation are from 8 to 18 years old.

comparisons to be valid. The following sections will discuss possible challenges to the interpretation of apparent time and will subsequently describe various methodological challenges of real-time analyses, as they relate to this present study.

2.5.1. Interpretation challenges of the apparent-time construct

It is possible to detect linguistic changes throughout the lifespan of an individual through panel studies in real time, namely a re-survey of the same group of informants after a period of time has elapsed (ex. Sankoff & Blondeau 2007, Wagner & Sankoff 2011). Changes in individual vernaculars may move either in the direction of an innovative features or in the opposite direction, and they will only end up influencing the rate at which the change will happen but not whether the change is actually taking place: interpretation of apparent-time results will therefore always under- or overestimate the rate of change, while correctly identifying the direction of change (Sankoff & Blondeau 2007: 582; Wagner & Sankoff 2011: 275).

In using, for example, both a panel study and a trend study, i.e. a re-survey of a community after a period of time has elapsed, Sankoff & Blondeau (2007) collected longitudinal data between 1971 and 1984 on the change in progress from apical [r] to dorsal [R] in 31 Montréal French speakers. They showed that whereas most individual vernaculars in their corpus were stable in adulthood, 9 speakers, mostly young adults, modified this feature in the direction of the change in progress (Sankoff & Blondeau 2007: 573). Even though the rate of change in individuals was not as stark as the one that occurred in the community as a whole, it appears that these late adopters played a role in accelerating the change in progress (cf. Boberg 2004[90]) (Sankoff & Blondeau 2007: 582). Sankoff (2006: 115) notes that had the

[90] Boberg (2004: 250) compared apparent-time phonological and lexical Montréal English data with real-time data from earlier studies of the same community in order to verify the assumptions of the apparent-time construct. His comparison revealed that some age-correlated lexical variables showed stability over speakers' lifetimes, suggesting ongoing change, and others showed changes in progress as well as change over speakers' lifetimes. Boberg (2004: 250) revealed that the mechanism of individual change was mostly the late adoption of

1971 data not been available and the 1984 data interpreted according to apparent time, she and Blondeau would have wrongly assumed that the speakers registering 90-100% of the innovative variant had all "begun their lives as children with those same values".

Moreover, Sankoff (2006: 114) compiled a list of 13 sociolinguistic studies from the 1960s and 1970s where a follow-up study was also conducted years later, since 1995, and she noted that in most panel studies, in the cases where there was a change in progress measured by a trend study, researchers found that "grouped data from the panel shows a modest increase in the direction of the change", and when researchers studied individual panellists "this result can typically be decomposed into a majority of speakers who remain quite stable, and a minority who change, often substantially". For Sankoff (2006: 114-115), panel studies therefore provide one clear result:

> as they age, people register lesser differences from their earlier selves than does the community over the same time interval, as measured by a trend study. This means first, that it must be younger speakers who are in the vanguard of change. Those adult speakers who change are (a) in the minority; (b) concentrated in the younger-age cohorts of adults and (c) make less significant advances than the community as a whole.

Another example can be found in Wagner & Sankoff (2011) who observed the rise of the periphrastic future (e.g. *je vais aller* 'I will go') at the expense of the inflected future (e.g. *j'irai* 'I will go') in Montréal French through the same panel study (1971-1984), but who recorded age grading in the opposite direction. As their informants had grown older, two-thirds of them had increased their use of the conservative (inflected) form: nonetheless these individual changes did not stop or reverse the linguistic change towards the periphrastic future, but only slowed its progress (Wagner & Sankoff 2011: 305). Evidence from Sankoff & Blondeau (2007) as well as Wagner & Sankoff (2011) therefore reveal clearly that change that is not the result

new variants by adults who learned older variants as children, rather than the rejection of new variants by older speakers, which is usually associated with the age-grading model. Most of the post-acquisition changes that were observed therefore contributed to accelerate rather than slow down changes in progress (Boberg 2004: 250).

of cyclical age-grading can take place across an individual's lifetime, whether in the direction of the direction of the innovative feature or not, and that it can also be compatible with an interpretation of change in progress in a community.

With regard to cyclical age-graded features, Labov (1963, 1994: 73) observed that their occurrence could be observed when younger speakers use seemingly innovative features which decline in frequency as these same speakers grow older but are adopted anew by the following generation of younger speakers. In other words, if older participants in an apparent-time study changed the way that they speak from the time of an earlier study "and if the earlier data suggest that they once spoke just as the younger participants do now, the apparent-time hypothesis must be rejected" (Boberg 2004: 251). As Boberg (2004: 257) explains:

> These changes are presumably conditioned by age-related shifts in social values and orientations, from innovation and counterculture in youth, to social ambition and increasing conformity in middle age, to conservatism in old age. The result of such instability at the individual level would be stability at the community level: no net change.

In the best-documented cases of age-grading, "the linguistic retrenchment occurs in adolescence and has the status of a coming-of-age ritual" (Chambers 2013: 310), but Chambers (2009: 201) notes that just a handful of cases of this kind have been reported in the literature. An example of this type of change is described in Macaulay's (1977) study of stigmatized glottal stops in Glasgow English: in apparent time, males belonging to the highest and lowest social classes exhibited stability, but the second-highest social class displayed a rise in the use of glottal stop in early adolescence, followed by a swift decline in adulthood (reported in Wagner 2012: 375). According to Chambers (2003) as well as Sankoff (2004), this behaviour observable in apparent time was an instance of age grading. Wagner (2012: 375) notes that further evidence from the behaviour of the girls and women in the sample supported this interpretation: unlike the boys and men, the girls and women in the higher social classes displayed a steady decrease in glottal stop use as they got older, without the peak

in adolescence that the boys had exhibited. Sankoff (2004) and Chambers (2003) both suggested that this peak represented a stronger pressure on the middle-class boys than on the girls to adjust their speech to covert/non-standard norms (of behaviour) during adolescence. They would have done so because "they shared assumptions about the social pressures experienced by men and women from different social classes at different stages of the life course" (Wagner 2012: 375). Relevant to these expectations is the concept of 'linguistic market', since (young) adults occasionally make sociolectal changes in response to the pressures of the marketplace (Wagner 2012: 375; Cukor-Avila & Bailey 2013: 253). For example, Sankoff & Laberge (1978) showed that market pressures, through the linguistic market index, were a significant factor in the incidence of three grammatical variables in Montréal French: alternation of auxiliaries *avoir* and *être*,[91] complementizer *ce que/qu'est-ce que*, and indefinite *on/ils*. Sankoff (2006: 112) observed different patterns depending on the nature of change:

> If a change is ongoing, older speakers as they age may change their speech, to some extent, in the direction of the change. In the case of sociolinguistic variables known to be stable, however, there may be a curvilinear pattern associated with age as well as with social class, whereby speakers in their mid-adult years, more implicated in the 'linguistic market' (Sankoff and Laberge, 1978) may show a greater use of standard variants than is typical of the oldest and youngest speakers.

In addition, Sankoff (2006: 112) points out that different phases of people's lives involve them to a lesser or greater extent in their relation to the standard language (Eckert 1997) and even also in their contacts with the opposite sex (Cameron 2000).

In his study comparing apparent-time phonological and lexical Montréal English data with real-time data from earlier studies, Boberg (2004: 266) found that the possibility of cyclical age grading was "not as serious an obstacle to apparent-time analyses of change in

[91] See more details on the linguistic market in §2.4.2.2.4. above. However, panel studies have revealed that patterns of auxiliary selection remain stable during a lifetime (Sankoff 2019), see sections §4.2.6. and §6.2.1.

progress as has previously been assumed". He observed instead that the patterns exhibited by older Montréal English speakers generally helped drive lexical changes to their completion, namely that the rate of innovative lexical forms in the community rose more rapidly than apparent-time evidence suggested it (Boberg 2004: 266). This type of late adoption would presumably be a result of the increased salience of innovative forms now employed by the majority of young people (Boberg 2004: 265). However, Boberg (2004: 265) found a more complex pattern with phonological variables whereby real-time evidence pointed to a change in progress but older speakers exhibited varying behaviours, with some rejecting and others adopting the changes. Interestingly, it was the oldest speakers who seemed to reject them, thus reverting to more conservative pronunciations in the subsequent study (Boberg 2004: 265). Boberg (2004: 265) hypothesized that:

> the likelihood of late adoption is inversely correlated with the degree to which a variable is structurally embedded: variables that are implicated in structural relations with other elements of the linguistic system, like phonemes and the contrasts they support, will be less susceptible to change in later life than variables like vocabulary items or matters of phonemic incidence, which typically bear no such structural relations and can be altered or replaced without systemic ramifications.

This result therefore supported the view that more abstract levels of grammar, such as phonology and syntax, are less likely to undergo post-acquisition change than the less abstract ones, such as the lexicon. For Boberg (2004: 266), this evidence of late adoption therefore challenges the view that apparent-time data are a less reliable alternative for real-time data, but such results could only surface when both types of evidence (apparent-time and real-time data) were taken into consideration. In the same line of reasoning, Sankoff (2006: 113-114) noted that, in her list of 13 sociolinguistic studies from the 1960s and 1970s where a follow-up study was also conducted years later, there are four possible outcomes, in the historical sense, revealed by the subsequent re-studies:

> First, if the original age distribution is repeated at the same level, we interpret the outcome as static age grading. Second, when we note a repeated age gradient but at a higher overall frequency of the change, we interpret the result as a real-time change. The third possibility is that all age groups display the same high level of the variable, which we interpret as the last phase of change going to completion. In this case the trend study should show no further increase on the part of a new generation of young speakers. Since eventually all changes are completed, it may be unreasonable to think that the absence of continuing change constitutes a failed prediction. The fourth possibility is that change is reversed, usually as the effect of stigmatization from above.

However, she observed that although one may find a combination of age grading and change in real time, in none of these follow-up studies the outcome was a result of age grading alone (Sankoff 2006: 114). The most important implication of this comparative finding for Sankoff (2006: 113) is that apparent time "is a truly powerful concept in locating the presence of change. In other words, a researcher who locates a gradient age distribution in a new community under study is virtually assured of having identified change, whether or not age grading is also involved".

So while individual vernaculars are generally stable during adult years, researchers have to be on the lookout for changes across the lifespan, be it in the direction of an innovation or towards the use of a more conservative variant. These changes simply make it more challenging to interpret the pace of a change, rather than invalidate the apparent time construct. And, as Chambers (2013: 310) points out, apparent-time surveys have overall the advantage of eliminating the years-long interval between surveys and allow data collection in identical settings and contexts, thus getting rid of comparability issues that can occur with real-time surveys. The following section will discuss these challenges.

2.5.2. Methodological challenges of real-time analyses

With real-time studies, researchers have two possible methodologies at their disposal. The first one is the comparison of new evidence with pre-existing data, which was a more common

approach for early researchers (Thibault 2001: 29; Sankoff 2006: 112). However, as Cukor-Avila & Bailey (2013: 254) rightly point out, in many cases "earlier linguistic evidence does not exist, and when it does, it often was not collected or organized in a manner that permits straightforward comparisons". For an example of this issue, Bailey *et al.* (1991) aimed to compare apparent-time distributions of fourteen features of Texas speech in the Phonological Survey of Texas (PST) and the Grammatical Investigation of Texas Speech (GRITS), both completed in 1989, with real-time evidence for those features from the Linguistic Atlas of the Gulf States (LAGS), data for which were gathered some 15-20 years earlier. However, the LAGS sample only included native Texans, while the PST and GRITS comprised a random sample of Texas residents including both natives and non-natives. Those differences between the sample populations of the LAGS and the PST yielded apparent-time distributions of monophthongal /ai/ in *night* in the PST suggesting stable variation in the use of this feature, while the differences between the LAGS and the PST suggested a change in progress (Bailey *et al.* 1991: 256). In addition to this sample issue, contradictory results were obtained for a grammatical variable, the use of multiple modal combination *might could* (roughly meaning "might be able"), when comparing apparent time and real time data because of methodological differences between two surveys: the LAGS relied on indirect elicitation to obtain tokens, while the GRITS relied on informants' self-reports on their use of the form (Bailey *et al.* 1991: 258). The use of apparent-time evidence would have therefore eliminated the variation caused by discrepancies in sample populations and in data collection methodologies when pre-existing data are compared to new ones.

The second option for performing real-time analyses is to re-survey the community (through a trend survey) or a group of the same informants (through a panel survey) after a period of time has elapsed, which can minimize the discrepancies just discussed when those factors are explicitly controlled for. However, as will be discussed below in more detail in §3.1.1., one of the main challenges of conducting trend surveys is rapid and on-going

demographic changes: the demographics of Montréal have been greatly transformed by municipal mergers, inter- and intraprovincial as well as international immigration, and the gentrification of underprivileged neighbourhoods, among other changes. Consequentially it was quite difficult to find not only 'native speakers' of Montréal French to participate in the present study but also speakers corresponding to social criteria used in the 1970s. During the creation of their Atlas of North American English, Labov *et al.* (2005: 27) also encountered such challenges when trying to find native speakers from Atlanta (Georgia), Dallas (Texas), and New York City. Typically, the more distant in time from the original survey the new survey is, the more different the sample populations are likely to be: the decline in rates of use of a certain variable could potentially simply reflect a change in demographics, more than a change in progress. Cukor-Avila & Bailey (2013: 255-256) hypothesize that if one had re-done their 1990 study of the use of *might could* in Texas in 2012, new data would have probably revealed a sharp decrease in the use of that feature overall. However that would likely occur because "the segment of the population that least uses *might could*, the Hispanics, is expanding rapidly, while the segments that use it the most (Anglos and African Americans) are either declining or remaining stable" (Cukor-Avila & Bailey 2013: 255).[92] Apparent-time evidence from a re-survey would also plausibly suggest that *might could* is disappearing due to the fact that, on average, the Hispanic population tends to be younger than the Anglophone one, resulting in younger age cohorts in future population samples having higher percentages of Hispanics and, most probably, fewer users of *might could* (Cukor-Avila & Bailey 2013: 256).

Sankoff (2006: 112) points out that even though trend studies are the only sure tool that can confirm change if the goal of the study is indeed to assess language change, panel studies are the "only way to discover how individual speakers of different ages are involved

[92] In 1990, Hispanics comprised 25,6% of the Texas population, while African Americans comprised 12,6% and Anglophones 60,6%. In 2000 and 2010, the percentage of Hispanics in the state rose to 32% and 37,6%, respectively, while the African-American population remained stable, at 12% in 2000 and 11,8% in 2010. The proportion of Anglophones declined to 52,4% in 2000 and to 45,3% in 2010. (Cukor-Avila & Bailey 2013: 255).

in linguistic change". Panel surveys, such as the ones on Montréal French (1971-1984-1995) presented in §2.4.2.1., are not directly affected by changes in the demography of a community, but do present the practical difficulties of, first, having to locate the same informants that were originally interviewed despite the mobility of many populations, especially in large-scale surveys, and, second, having to interview them again in the same way (Cukor-Avila & Bailey 2013: 257). As a consequence of this first difficulty, it is therefore challenging to re-interview all of the original participants in a panel study, which will inevitably yield a smaller and/or less representative sample for the re-survey. As to the second challenge, Schilling-Estes (2005: 228) points out that in theory "in order to obtain comparably natural, vernacular data and hence a truer picture of [an informant's] speech over time, we need to locate the same interviewer as in [the original survey] and have her interview [the informant] in the same way, in the same location". But not only is this methodology highly impractical, for Schilling-Estes (2005: 229) it is also meaningless because role relationships change constantly, even in the course of a single conversation (for example when an interviewee starts asking questions of the interviewer). So even if a participant were to be interviewed years later by the same interviewer from the original interview, "in the same setting, there is no guarantee that the two would stand in the same relation to one another and hence produce comparable speech data" (Schilling-Estes 2005: 229).

In the best-case scenario, however, researchers are able to compare apparent-time evidence to real-time one by carrying out "replications of a number of previous sociolinguistic and dialect surveys, adding a real-time component to the study of language change and variation" (Sankoff 2006: 112). Despite the difficulties of obtaining real time data that are comparable to newly collected ones, it is in these combined conditions that "researchers will attain the kind of insights that revolutionize our understanding of language change and make the study of change in progress the apogee not only of sociolinguistics but perhaps, as Chambers (2009: 160) says, of contemporary linguistics" (Cukor-Avila & Bailey 2013: 259).

The present study will make use of both types of analysis, by being first being conducted in apparent time and by then using previous studies on Montréal French auxiliary alternation for a real-time comparison. This methodology will be discussed further in §3.5.1.

2.6. Discussion and research questions

It appears that speakers of various Romance varieties try to make sense of the distribution between HAVE and BE by mapping it into formal distinctions,[93] as seen in §2.2., but as noted earlier that neither the the semantic nor the syntactic frameworks of split auxiliary selection can fully account for the variation that exists in the French auxiliation system.

Proponents of formalizing semantic theories, such as the Auxiliary Selection Hierarchy (outlined in §2.2.2), in addition to analysing only standard (usually written) languages[94] and hence reducing language to arbitrary systems rather than analysing 'real language in use' and 'non-standard' varieties, treat the languages under consideration as stable synchronic entities. The ASH does not take into account how much variation there was before the implementation of codification processes. Moreover, this semantically-driven approach does not explain how different auxiliaries can be used with the same verb, in the active voice, in the same language, and by the same speaker, as can be observed in contemporary speech in many varieties of French (see §2.4.2.). It also leaves no place for sociolinguistic considerations (intra- and inter-speaker variation in terms of age, gender, socioeconomic

[93] In psycholinguistic terms, it could be argued that some of the patterns described earlier can be explained through the concept of synonymy avoidance or by the Principle of Contrast theorized by Clark (1987, 1993: 64), which states that "speakers take every difference in form to mark a difference in meaning". Regarding the pronominal verbs, Chaudenson *et al.* (1993: 24) explain that, on top of the diachronic explanation, the use of *avoir* with pronominal structures also has a *motivation intrasystémique* 'intrasystemic motivation' insofar as it can be observed in situations of language acquisition and language learning. In fact, Heinen & Kadow (1990: 65) noticed that children acquire *avoir* as perfect auxiliary before they acquire *être* as such. Hallion (2000: 355) also confirms that French children (3- to 4-year-olds) sometimes produce the following pronominal constructions with *avoir*: *Je* ***m'ai*** *fait piquer* 'I have been stung' and *Regarde ce que je* ***m'ai*** *fait* 'Look what I have done to myself'.

[94] For example, some of Giancarli's (2011) French data come from a corpus of French plays, as stated in §2.4.2.2.6.

background, etc.) and linguistic constraints (apart from lexical effect), when numerous studies (see §2.2.1.1. for Romance examples and §2.4.2. for French) have shown that auxiliary selection can vary along those lines.

As to the syntactic analyses mentioned in §2.2.1., while they rigorously cover regional variation in an exhaustive number of non-standard (Italo- and Ibero-)Romance dialects, there is little or no mention of how the data for each of the studies quoted was collected. It also seems highly improbable that there would only be one Romance variety, i.e. Canadian French (exemplified only by Sankoff & Thibault 1977; King & Nadasdi 2005; Rea 2014), showing social, otherwise "free", variation in its auxiliation.[95] As Tristram (2014: 6) points out:

> Even if phonological and morphosyntactic variation turn out to pattern in significantly different ways (and there is some suggestion that this may be the case; cf. Armstrong 2001), it seems almost inconceivable, given what is now known about phonological variation and change, that social factors would not play some role in morphosyntactic variation and change as well.

Since free variation is extremely rare, and given the various auxiliary distributions summarized in §2.2.1., why would Canadian French pattern so differently from the other Romance languages? I would argue that the reported discrepancies between auxiliation patterns in varieties of Canadian French and in other Romance languages stem mainly from a difference in data collection methodology because very little metadata is collected and analysed in these dialectology studies. This suggests that Romance auxiliation data therefore offer a fertile testing ground for applying a variationist sociolinguistic methodology.

It has been suggested by Thibault & Sankoff (1997, not paginated) that in Montréal:

[95] It should be noted that Ledgeway (2000: 185-186) did find different auxiliation patterns in Literary, Urban, and Peripheral Neapolitan. Literary Neapolitan has a system similar to standard Italian, where argument structure determines auxiliary choice, Urban Neapolitan generalizes HAVE, and peripheral varieties (Torre del Greco, Torre Annunziata, Pompei, Sorrento) show a person-based pattern where BE appears with first and second person subjects and HAVE with third person ones.

> although there is a tendency for the non-standard use of *avoir* to be associated with working class speakers, this applies differentially to the various verbs, and the social class based pattern is far from categorical. A semantically driven association between auxiliary use and verb meaning is weakly motivated at best. Rather, there seems to be a lexically-focused distribution.

This would imply that the variation in Montréal in 1971 and 1984 was lexically arbitrary, but no non-French Romance languages seem to display this pattern. It might also be the case that the newly collected data show that the variation is after all simply constrained by semantic/grammatical factors, like the other Romance languages.

This chapter has also discussed the various variables, both external/social and internal/linguistic, that have been shown to influence *avoir* use in previous studies of North American varieties of French. What emerges from this survey is that patterns of auxiliary alternation in French are much more complex than what purely syntactic or semantic approaches are able to capture.

Further real-time and apparent-time analyses are required to test whether the same variables of influence on auxiliary alternation (both social and linguistic) in Romance languages, and French in particular, are still at play in newly collected data and whether additional ones can be uncovered. A variationist methodology is used in my study in order to collect data as close to "real language in use" as possible, which helps us fill the gap about how auxiliary selection behaves in Montréal French today. Building on the conclusions of Sankoff & Thibault (1977) and Thibault & Sankoff (1997), I want to examine the state of this alternation today given the great sociodemographic changes that have taken place in Montréal in the last 50 years.

Since 1971, the province of Québec has witnessed an important wave of nationalism[96] closely associated with the officialization of French as the only official language of Québec, with the promulgation in 1977 of the quite controversial Charter of the French language (also

[96] The province of Québec had two referenda on the question of sovereignty and independence, one in 1980 and one in 1995. Both times, the proposal to pursue secession was rejected by a very small margin.

known as *Loi 101*). Adopted in a bid to protect the French language while freeing the province, and especially the city of Montréal, from the dominance of English, the Charter paved the way for the francization of the province's government, courts, businesses, workplaces, and education system.[97] This led to a mass exodus of anglophones (and allophones[98]) towards other Canadian provinces. While the Charter has been amended several times over the years, its most controversial articles still concern the obligatory use of French for all commercial signs and publications, as well as including severe restrictions on enrolment into anglophone schools. It is more than likely that without such measures, Montréal would be predominantly an English-speaking city today. The linguistic landscape of Montréal, as well as the relationship between its francophone speech community and its language have been greatly transformed since 1971. And that is why a follow-up on auxiliary alternation is therefore needed to determine: 1) the state of auxiliary alternation in Montréal French today compared to 1971 (as well as to 1984 and 1995), and in which direction the changes have been observed; 2) whether there has been a change in the social and linguistic distribution and conditioning of the variable, and if so, what form it has taken; 3) whether the intransitive and pronominal data will turn out to be distributed in similar social and linguistic patterns to those that have already been recorded in other French and Romance varieties; and 4) what would be the implications of such findings.

This thesis also investigates in a rather exploratory way auxiliary variation in pronominal verbs, excluded by Sankoff & Thibault (1977),[99] since there is no doubt that Montréal French also presents certain irregularities in this regard (e.g. *Je m'ai fait mal* 'I (have)

[97] "Introduced by [Québec MP] Camille Laurin [one of the early founders of the Québec sovereignty movement], Bill 101, *Charte de la langue française* (1977), made French the official language of government and of the courts in the province of Québec, as well as making it the normal and habitual language of the workplace, of instruction, of communications, of commerce and of business."
(http://www.thecanadianencyclopedia.ca/en/article/bill-101/) [webpage accessed March 26, 2017]

[98] These were mostly immigrants who had hoped to assimilate to English culture and/or wished to send their children to English school.

[99] Sankoff & Thibault (1977) had excluded from their research this verb category because they assumed it to be more conservative in its auxiliary selection, even though they did not give a diachronic perspective against which to evaluate that.

hurt myself'). This thesis also tests whether it is possible to confirm my findings via two triangulation methods: the analysis of grammaticality judgement data from my 48 participants and self-reporting judgements from the survey *Français de nos régions* (Avanzi *et al.* 2016).

Chapter 3. Methodology

This chapter covers the methodology employed 1) to build an original sociolinguistic corpus of Montréal French; 2) to operationalize the dependent variable and the various social and linguistic independent variables that are tested for possible influence on auxiliary alternation; 3) to circumscribe the variable context and transcribe the relevant data; 4) to establish exclusion criteria for exceptional tokens; 5) to analyse the production data collected during the interviews; and 6) to analyse the grammaticality judgement data from my fieldwork and from the corpus of self-reported judgements *Français de nos régions* (Avanzi *et al.* 2016), with which my trend study findings are triangulated.

3.1. Constituting the corpus

3.1.1. The definition of "Montréal French"

In recent years, various linguists studying Montréal French have focused on individual neighbourhoods rather than studying Montréal as a single speech community. As part of the umbrella project led by France Martineau, *Le français à la mesure d'un continent : un patrimoine en partage*, scholars such as Blondeau, Martineau, Tremblay, & Frenette (2012) as well as Blondeau, Martineau, & Tremblay (2013) have studied Montréal French not by covering the whole city but by concentrating on individual primarily French-speaking neighbourhoods – in particular Hochelaga-Maisonneuve (2012), which is slowly becoming gentrified, and Saint-Michel-Montréal-Nord (2013). The Hochelaga-Maisonneuve corpus (2012) includes 48 semi-structured interviews with participants aged between 18 and 89 years old.[100] As to the Saint-Michel-Montréal-Nord corpus (2013), it includes semi-structured interviews with participants

[100] At the time of writing, only 18 interviews had been made available and only two of these had been conducted with speakers below the age of 55.

aged between 19 and 54 years old who have been raised in a multicultural neighbourhood and who are themselves, in majority, of immigrant background. The latter corpus aims to document the contribution of young and Neo-Montréalers on the dynamics of spoken Montréal French. A third Montréal neighbourhood corpus, of the Ahuntsic-Cartierville neighbourhood, has also been compiled by Bigot & Papen (2018).[101]

Since the aim of the present project is to reproduce as closely as possible the study of Sankoff & Thibault (1977), Montréal is studied here in its entirety so that my results can be usefully compared with theirs. By re-studying a speech community through the creation of a new sample with new speakers, including those from younger generations, and matching the earlier sample(s), outlined in §2.4.2.1., as closely as possible (in terms of age groups, social status, etc.), this sampling choice allows me to treat this project as a trend study (see §3.5.1.).

At first, determining whom to include in the sample seemed very straightforward: the speakers would have to fit the same criteria (place of birth/childhood, current *arrondissement* 'neighbourhood'/municipality where they are living, etc.) as the ones who were interviewed for the Sankoff-Cedergren corpus of 1971. However, as Schilling (2013: 56) puts it "we may have trouble locating new speakers who fit the original social criteria (and indeed have to question whether such speakers would now yield the most accurate picture of language use in the now-changed community)".

In fact, when trying to find speakers who fit the minimum requirement (namely, that of being a native speaker of Montréal French), it was quite difficult – despite being a native speaker myself – to find 'old stock' Montréalers: people who were born in Montréal, were raised in Montréal by Montréalers, and who had spent all their life there as well. Consistent with Schilling's warning, the demographics of Montréal have changed considerably in 50 years (Lamarre 2016): the metropolitan area of Montréal now includes many more towns than in

[101] Very few details have yet been made available regarding the creation and the composition of that corpus.

1971 because of municipality mergers, and it is now common to speak of the Greater Montréal area, encompassing municipalities surrounding the Island of Montréal.[102] Moreover, the population of the Island of Montréal amounts to approximately 2 million inhabitants and a third of them are immigrants; 56% of Montréalers were either born in a different country or have at least one parent who was born outside Canada (*Profil socio-démographique de la ville de Montréal* 2014). Rather less than half of the population of Montréal (i.e. slightly fewer than 1 million people) now have French as a first language (*Profil socio-démographique de la ville de Montréal* 2014). It is also important to bear in mind that the numbers do not even distinguish internal migrations within the province of Québec: many Quebecers come to Montréal in order to study, work, etc. and end up settling there permanently (Institut de la statistique du Québec 2000).[103] Many Montréal working-class neighbourhoods also undergo gentrification (Marceau 2016; Montpetit 2016; Gagnon 2016; Corriveau 2016; Hays 2016) and traditionally French-speaking neighbourhoods have now become multicultural and multilingual (Radio-Canada 2012). The French/English divide is considerably more blurred than in the 1970s, even geographically (Boychuk 2017; Presnukhina 2016; Paillé 2011; Statistique Canada 2007).

Therefore, while the label "native speakers of Montréal French" might have been a fairly straightforward and self-explanatory one in 1971, it is certainly not the case anymore. Labov *et al.* (2005: 27) encountered the same problem when creating their Atlas of North American English, especially when trying to find native speakers from Atlanta (Georgia), Dallas (Texas), and New York City. Labov *et al.* (2005: 27) recognize that this issue can seem to invalidate the goal of the study:

[102] http://www.toponymie.gouv.qc.ca/ct/ToposWeb/Fiche.aspx?no_seq=42164 [webpage accessed April 11, 2017]

[103] In 1991-1998, intraprovincial immigrants were the second largest source of demographic growth in Montréal, after births (Institut de la statistique du Québec 2000: 26).

> It may seem paradoxical that it is difficult to locate speakers with the desired characteristics when the goal is to represent the speech patterns of the community as a whole. But it is not uncommon to find that the main stream of vernacular tradition is obscured by the presence of large numbers of recent arrivals in the adult population.

Though this makes a good case for focussing on a single neighbourhood and treating it as a microcosm of Montréal, as did for example Blondeau, Martineau, Tremblay, & Frenette (2012)[104] and Blondeau, Martineau, & Tremblay (2013), this study cannot rely on such methods because no neighbourhood would really be representative of the entire Montréal situation.

This is also why I include speakers from all over the *Grand Montréal*/Greater Montréal area[105] and native Québécois French speakers who might not have been born and raised in Montréal, but come from neighbouring regions, have lived in Montréal for most of their economic life and are now established in the *Grand Montréal,* working or studying full-time on the Island of Montréal. These inclusions reflect more accurately the linguistic situation of Montréal.

3.1.2. Judgement sampling

As stated in §2.4.2.1., the original Montréal survey of 1971 realized by Sankoff & Cedergren made use of a random stratified sampling method to obtain a sample of 120 speakers, by first identifying Montréal neighbourhoods that were mainly French-speaking and then ensuring that the participants selected from these areas were native French speakers who had resided in the city at least since the age when they began primary school. Interviewers went door-to-

[104] For comparison purposes, the researchers involved in the constitution of that corpus believed that the neighbourhood Hochelaga-Maisonneuve was particularly representative, in sociolinguistic terms, of how Montréal was in 1971 (Blondeau & Tremblay, personal communication 2017).

[105] It includes an inner ring composed of densely populated municipalities located in close proximity to Downtown Montréal (i.e. the Island of Montréal, Laval, and the Urban Agglomeration of Longueuil). The outer ring is composed of low-density municipalities located on the fringe of Metropolitan Montréal. For a detailed list of all municipalities included, see http://cmm.qc.ca/fr/a-propos/municipalites/ [webpage accessed April 12, 2017]

door in various French-speaking neighbourhoods of different socioeconomic levels to see whether anyone wished to participate in their study of Montréal French. The speaker sample was also stratified internally according to age (four groups of 30 speakers each), sex (each age group contained 15 male and 15 female speakers), income level, and geographical region.

While this project aims at emulating the methodology of the 1977 study as closely as possible, time constraints made it impractical to go door-to-door in order to find participants to obtain a balanced sample.[106] I therefore work with judgement sampling, targeting "participants that meet the predetermined criteria of the study, such as relevant social categories […] [and making] use of extended social networks of the researcher and the researcher's contacts within the community" (Hoffman 2014: 31), by finding participants through personal acquaintances. Judgement sampling is consequently not random and does not "ensure statistical representativeness" (Schilling 2013: 35), but as Schilling also points out: "we can usually be confident that the insights we obtain are generalizable to the larger population thanks to the relative uniformity of linguistic vs other social behaviors." Judgement sampling, which is therefore a balance between random sampling and the social network approach (Tagliamonte 2006: 27-28), is the most common fieldwork technique for both methodological and pragmatic reasons (Hoffman 2014: 31; Tagliamonte 2006: 27-28) and has become the "consensus in the field" (Tagliamonte 2006: 27).

The fact that the speakers knew me personally (or knew of me) and that I belonged to this speech community made it easier during the sociolinguistic interviews to gain access to the vernacular, also called "informal speech", "everyday speech" (Sankoff 1974, 1980a: 54), "real language in use" (Milroy 1992: 66), or "spontaneous speech reserved for intimate or casual situations" (Poplack 1993: 252). This form of speech is thought to be the most systematic (Tagliamonte 2006: 8), since it is learned unconsciously and based on internalized

[106] There are also risks to the investigator's personal safety.

rules, as opposed to formal speech or written language, which instead tends to reflect explicitly learned prescriptive rules. Knowing the informants or knowing of them provided me with an additional advantage: it was possible to place them roughly in the various social categories (gender, age, and socioprofessional status) before obtaining all the personal information needed to assign them to a cell with more certainty.

To ensure that the study be statistically valid, a minimum of four or five speakers in each cell was needed (Hoffman 2014: 30). Using a minimum of five speakers per cell with three age groups and three socioprofessional statuses (SPS) would have produced a very high number of speakers to interview.[107] This meant that either the minimum quantity of speakers per cell or the quantity of subcategories of age/socioprofessional status had to be reduced. It was decided not to study the oldest section of the population, over the age of 66 years old in 2016 (which would overlap with Sankoff & Thibault's young to middle age group in 1971), in order to maximize the number of new speakers available.[108] This decision slightly limits the parameters of an apparent-time analysis (see §2.5.) because even though no age effect was detected (see Tables 4.2. and 4.5), the inclusion of an older age group might have revealed one since the overall rate of *avoir* has drastically diminished between 1971 and 2016 (see Table 4.1). However, this sampling decision does not necessarily deny me the possibility of discovering whether the oldest age section of the population has changed linguistically because recent research by Sankoff (2019: 198) suggests that speakers do not change their auxiliary selection patterns across their lifespan. This will be discussed further in §4.2.6. and §6.2.1.

It should be pointed out that new statistical tools, such as *Shiny Rbrul* (Johnson 2017, see §3.5.2.), allow the user to study a continuous factor as such. While this does not mitigate

[107] 5 (speakers) x 2 (genders) x 3 (age groups: 18-35 y-o, 36-65 y-o, and 66-85 y-o) x 3 (socioprofessional status: LOW, MID, HIGH) = 90 speakers

[108] In hindsight, since MID and HIGH speakers tended to behave in the same way with regard to their influence on the data (see Models 2 and 3, respectively in §4.2.2.2. and 4.2.2.3.), it would have been possible to reduce the number of SPS levels to two and to rather include the oldest section of the population instead.

the absence of older speakers, a statistical calculation with a continuous factor provides a more fine-grained analysis that can detect a potentially significant effect on the variation based on data from each speaker of the corpus, that is for each increasing or decreasing year of age, instead of simply labelling speakers as 'young' and 'old' or as belonging to certain age groups (see §3.2.2.3. for more details). Rather than comparing two data points, namely the average *avoir* rate of all the speakers belonging to the 'young' category vs that of all 'old' speakers, each individual age is included in the model, with 44 years separating the youngest 2016 speaker from the oldest one.

A new sampling grid with two age groups would therefore include 60 speakers,[109] which would still impose severe impracticalities. A minimum of four speakers in each cell, i.e. one or two more speakers per age/sex/SES category combination than in the 1971 corpus (see §2.4.2.1.), would then also allow the maintenance of the three basic social macro-categories while yielding a total number of speakers that could practically be handled: it was decided to limit this project to a survey of 48 participants. The distribution of the speaker sample is presented in Table 3.1.

	Women		**Men**		
	18-35 y-o	**36-65 y-o**	**18-35 y-o**	**36-65 y-o**	
SPS LOW	4	4	4	4	
SPS MID	4	4	4	4	
SPS HIGH	4	4	4	4	
Subtotals	**12**	**12**	**12**	**12**	**Total: 48**

Table 3.1 Distribution of the speaker sample

[109] 5 (speakers) x 2 (genders) x 2 (age groups: 18-35 y-o and 36-65 y-o) x 3 (socioprofessional status: LOW, MID, HIGH) = 60 speakers

While there could be a great difference in auxiliary use between 36-year-olds and 65-year-olds, the cut-off age of 35 was selected based on the 1971, 1984, and 1995 Montréal corpora sampling grid. The decisions regarding how the SPS categories were assigned will be discussed in detail with reference to the coding of independent variables in §3.2.2.

I initially found speakers through personal contacts, and then by snowball-effect (Milroy & Gordon 2003: 32) – asking the participants whether they would be able to introduce me to various acquaintances of theirs that could fit certain specific social descriptors. The cells were all successfully filled according to the sampling grid presented in Table 3.1.

The sociodemographic profile of the 48 participants is presented in Table 3.2. In order to preserve the anonymity of my participants, each speaker was given a pseudonym starting with the same letter as their real given name. Other identifying information (such as names of family members, house addresses, places of employement, etc.) mentioned during the interview were redacted or modified in the transcriptions.

Speaker pseudonym	Gender	Age	SPS level (LOW, MID, HIGH)	Occupation	Place of residence (island vs suburb)
Alexis	M	26	M	VFX[110] artist	Island
Amélie	F	22	H	Masters student	Island
Annie	F	27	M	Translator	Suburb
Annouck	F	52	M	Elementary school teacher	Island
Carl	M	35	L	Internet service technician	Suburb
Caroline	F	27	L	Pastry cook	Island
Charles-Antoine	M	29	H	Lawyer	Island
Christine	F	40	H	Soundwoman	Island
David	M	28	H	Masters student	Island
Denis	M	49	M	Crane operator/docker	Island
Denise	F	58	M	Retired nurse	Suburb
Dominic	M	22	L	Truck driver	Suburb
Florence	F	37	H	Business associate/ Entrepreneur in the consulting sector	Island
Francis	M	53	L	Unemployed/ writer	Island
Gaëtan	M	61	H	Engineer	Island
Guylaine	F	53	H	Nurse	Suburb
Hugo	M	18	L	CÉGEP[111] student	Suburb
Jacynthe	F	50	L	Bartender	Suburb
Jean-François	M	26	H	Lawyer	Island
Jessica	F	25	M	Occupational therapist	Suburb
Joël	M	26	L	Storehouse clerk	Suburb
Johanne	F	45	M	IT manager	Suburb
Julie	F	28	L	Unemployed/ CÉGEP student	Suburb
Justin	M	26	H	Financial advisor	Island
Kim	F	29	L	Hairdresser	Suburb
Linda	F	49	L	Housekeeper	Suburb
Madeleine	F	55	H	Retired financial analyst	Island

[110] "Visual effects".

[111] *Collège d'enseignement général et professionnel* (CÉGEP): Québécois students who wish to pursue post-secondary education in a Québec university must attend a 2-year college (CÉGEP) before enrolling. Before beginning post-secondary studies, Québécois students complete one grade fewer than other North American students, finishing high school at grade 11 instead of 12. CÉGEP can also prepare students for a technical profession (3-year program). CÉGEP resembles closely higher school leaving examinations. See http://www.cegepsquebec.ca/en/ [webpage accessed on March 29, 2017]

Marc	M	39	M	Secondary school teacher	Suburb
Marie-Jeanne	F	22	H	Undergraduate student	Island
Marie-Laurence	F	25	L	Housekeeper	Island
Marie-Lou	F	24	M	Undergraduate student	Suburb
Mario	M	43	M	Photographer	Island
Martin	M	45	H	Entrepreneur in the event sector	Island
Martine	F	55	L	Transcriber	Island
Mathieu	M	26	M	Sales' coordinator	Island
Maxime	M	25	M	Doctoral student	Suburb
Nathan	M	18	M	CÉGEP student	Island
Paul	M	40	H	CÉGEP teacher	Suburb
Philippe	M	60	L	Waiter	Island
Rachel	F	42	M	Support worker/ stay-at-home mother	Island
Richard	M	53	H	Director of photography	Island
Sabrina	F	28	H	Lawyer	Island
Sara	F	37	L	Daycare teacher	Suburb
Sophie	F	26	M	Call centre coordinator/ Undergraduate student	Island
Steve	M	42	M	Engineer	Suburb
Sylvain	M	39	L	Welder	Suburb
Virginie	F	28	H	Lawyer	Island
Yves	M	59	L	Construction manager	Suburb

Table 3.2 Sociodemographic profile of the 48 participants

3.1.3. The fieldwork and the sociolinguistic interview

My fieldwork was spread over two visits to Montréal in order to be able to identify gaps in the sample, and select speakers before a second visit, in order to fill those cells. The first visit took place in March and April 2016, when 43 speakers were interviewed. After determining how to calculate with more precision the socioprofessional status index (see §3.2.2.4. below), I realized that there was some overlap in the speakers' metadata and had interviewed eight speakers of the same age group/socioprofessional status too many. I went back to Montréal in August 2016 in order to interview the missing thirteen speakers so as to yield a balanced

sample of 48 speakers, as presented in Table 3.1. The interviews took place at the speakers' home or at their workplace. In a few cases, the interviews took place at the home of a common acquaintance.

The data-collection interview was divided into six parts: having the speakers read first the participant information sheet again[112] (see Appendix 3A.2), having them sign the consent form (see Appendix 3A.3), collecting the speaker's personal information (see Appendix 3B), conducting the sociolinguistic interview, completing the grammaticality judgement task (see Appendix 3C), which will be described in more detail in §3.6.1 and analysed in Chapter 5 below, and finally administering a linguistic attitudes questionnaire[113] (see Appendix 3D).

It was thought best to tell the participants that the goal of the study was to observe how native Montréal French speakers remember events that have happened in the recent past, rather than telling them that it was simply about Montréal French.[114] In a pilot study (Rea 2014), when speakers were invited to take part in a study about Montréal French, most were reluctant to be interviewed/recorded because they thought that they did not speak "well" enough to be good informants.[115]

The sociolinguistic interviews lasted at least one hour per person and, when possible, approximately half of them were conducted with speakers in pairs in order to facilitate the fluidity of the exchange. Labov (1972a) used group interviews in his early studies of African-

[112] The participants were initially given the participant information sheet when they were first contacted and invited to participate in the study.

[113] Labov's (1963) study of Martha's Vineyard confirmed that the speakers' positive attitudes towards the community made them adopt the local variant more so than the others. Kraus (2006) carried out a study with 126 Québécois students showing that language attitudes towards Québécois French are indeed complex and split, even within a single speaker: a majority of participants agreed that le *franco-québécois* (a term referring to the colloquial form of Québécois French) had the same worth as the standard, while most of their other answers indicated a preference for le *français québécois standard*. With that in mind, I administered a linguistic attitudes questionnaire in order to observe whether positive attitudes towards Québécois French and culture would correlate with *avoir* selection. However, I subsequently decided that the analysis of linguistic attitudes would be best suited for a separate research project.

[114] This strategy was included in the approval by CUREC 1A (see Appendix 3A.1).

[115] Linguistic insecurity – the feeling of anxiety, self-consciousness, or lack of confidence in the mind of a speaker surrounding their use of their native language – is prevalent all over French Canada. For works that specifically discuss linguistic insecurity in Québec, see Remysen (2018, 2004, 2001), Bernard Barbeau (2017), Beaudoin-Bégin (2015), Kircher (2012), and Cajolet-Laganière & Martel (1993), among others.

American teenagers in Harlem and found that interviewer effects were minimized. Similarly, the interviews from the Sankoff-Cedergren corpus of 1971 were not always one-to-one either, since the interviewers encouraged family members, spouses or parents, to participate during the interviews (see §2.4.2.1.). The people being interviewed in my groups therefore knew each other very well, and it was expected that they would talk more with one another than with the interviewer, which would obviate some of the artificiality introduced by the interview event. This was indeed the case in my study. Some examples of relationships in the groupings included life partners, parent and child,[116] best friends, siblings, etc.

The audio recordings were done at a sampling rate of 48 kHz with a PMD660 Marantz recorder and two Audio Technica AT803b lapel microphones connected to the recording device on two different channels, one for the participant and one for the interviewer (or one for each of the participants if they were interviewed in a pair).[117] The audio files were saved on a Kingston Technology 4GB compact SD memory card before being transferred onto an external hard drive, and the files were deleted from the memory card as soon as they were transferred. The protocols followed for ensuring proper data storage and protection are outlined in Appendix 3A.2.

Given the nature of the variable under study, the sociolinguistic interviews had to be somewhat different from the traditional Labovian interview, which works best when the goal

[116] It did not take very long to realize that the pair 'parent and child' was not suitable to elicit a maximum amount of spontaneous speech, because it seemed to be frequently the case that the child (or the parent) felt that they could not speak completely freely.

[117] While the equipment used to record the sociolinguistic interviews was ideal to produce high quality audio files suitable for future phonetic analyses, trading such equipment for a more discreet kind might have put the speakers more at ease. Since every speaker had a lapel microphone connected to a battery pack which was itself connected to the recording device, they could not move freely during the interview. Since most speakers were already put off by having to sign a consent form prior to the start of the interview, the fact that they had to wear a lapel microphone made the whole process very formal. Less intrusive equipment, such as a smartphone, might have allowed the speakers to forget more easily that they were being recorded, yielding a more informal speech register, and thus probably more *avoir* tokens. This hypothesis stems from anecdotal evidence observed during my fieldwork: mid-way through my interview with speaker Johanne and her sister, the memory card of my recording machine stopped working. After I replaced it with a new one, the two participants suggested starting the interview again, answering my questions a second time. I noticed that when they 'answered' the same questions again, the two speakers used far fewer *avoir* tokens the second time around, especially with *passer (proche)* and *tomber*. They were trying to remember how they had first recounted the event, and not the event itself.

is to collect phonological variants. Rather than encouraging the participants to speak about more or less anything for about an hour, which would not guarantee a sufficient number of tokens, the interview included questions specifically designed to elicit a maximum number of periphrastic tense tokens of the 16 verbs studied. Various strategies, some of them tested in the pilot study (Rea 2014), were developed in order to ensure the collection of a sufficient quantity of tokens. Professor Raymond Mougeon of York University in Toronto, who had worked on the same phenomenon in Ontarian French, suggested the use of questions about travels, holidays, and house moves because such questions had worked well for him in the past. Some questions were also inspired by sociolinguistic interview questions from the previous Montréal corpora (see §2.4.2.1.). Here are the questions[118] that were asked (usually in this order, which seemed to flow more naturally) to the participants in *tutoiement*:

1. How did the two of you meet? (when the speakers were in a pair)
2. How did you become an X or how/why have you started to become interested in X?
3. Could you tell me about a memorable trip that you have made or a favourite holiday that you have taken?
4. How did your last (house) move go?
5. What kind of renovation work have you done in your house?
6. How do you remember your first love?
7. How do you remember the 1998 ice storm?
8. What kind of memories do you have about being taught how to ride a bike or about learning how to swim? Or learning how to drive?
9. What is the most serious accident that has ever happened to you?
10. Could you tell me what you did yesterday, with as many details as you can remember?

[118] The questions were asked in French.

11. Have you ever been in a linguistic conflict? In a situation where you could not be understood by English speakers or by speakers of European French? What happened?

It goes without saying that if a speaker mentioned something related to one of the themes stated above while answering another question (for example: the mention of an accident during the 1998 ice storm), I would then take the opportunity to ask them about it as a follow-up question. This technique allowed me to ask most of the interview questions without interrupting the flow of the conversation.

Question 9 is a slightly modified version of Labov's "Danger of Death" question: "Have you ever been in a situation where you nearly lost your life? When you thought this is *it*?" (Labov 1972b: 92). In the pilot study (Rea 2014), the Danger of Death question was used more or less word-for-word and it would sometimes elicit answers related to severe illnesses. This was problematic because it had the effect of creating a general uneasiness in the room and seldom yielded tokens of the verbs under study. For these reasons, the original question was modified. The objective of inserting a Danger-of-Death-like question in the interview was two-fold. The most obvious reason had to do with lowering attention to speech and reducing the shyness experienced by some participants more reluctant to express themselves when recorded or when talking to a stranger, as explained by Schilling (2013: 102): "…not only are interviewees engrossed in their harrowing stories, but they feel compelled to convince their audience that the situation really was quite serious, and their concern for producing gripping stories outweighs any reticence they may feel about talking freely in the interview situation." In addition, the modified question allowed me to collect a fair quantity of relevant auxiliary tokens, since the narrative of accidents often correlated with the utterance of motion verbs (most of them belonging to the list of the 20-or-so verbs studied in this research) in periphrastic past tenses. One apparent disadvantage of the use of this question, however, was

the not infrequent use of historic present to render the story even more real to the hearer. The use of historic present will be looked at in detail in §4.4.2.

3.2. Operationalization of the variables

This section describes the various variables (internal/linguistic and external/sociolinguistic) that could affect the auxiliary variation in Montréal French. It also explains why each of these variables has been selected. The approach to the sociolinguistic variables is quantitative, unless stated otherwise, with a new version (*Shiny Rbrul*, Johnson 2017) of the statistical software *Rbrul*, an add-on package in *R* (*R* Core Team 2004-2019) specifically tailored for sociolinguistics and developed by Johnson (2009), which ranks independent variables (or factor groups) according to their impact on the data.

3.2.1. Dependant variables

The dependent variable tokens were first coded as either being intransitive or pronominal verbs before being coded as surfacing with *avoir* or *être*. Some tokens appeared with both auxiliaries in the form of false starts with *avoir* followed by self-corrections with *être*[119] and were also coded separately as 'false starts'. Additionally, two potential avoidance mechanisms were coded separately, namely uses of *avoir été* 'to have been' for *être allé* 'to have gone' and the use of historic present instead of *passé composé*.

3.2.1.1. False starts with avoir *followed by self-corrections to* être

False starts with *avoir* followed by self-corrections with *être* were excluded from the main analysis because they vary a lot depending on how much and what kind of material appears

[119] The false starts were always in this order.

in the false start itself and in the self-correction to *être* that follows. It is therefore impossible to know with certainty whether the speaker was planning on using the same verb. Eight different combinations of false starts and self-corrections were recorded in the 2016 corpus. The first one simply repeats the subject and the auxiliary, changing it to *être*, as in (20).

(20) *Faque là* ***on a, on est allés*** *en excursion dans une autre ville de la Colombie-Britannique.* (Justin)
'So then we, we went on an excursion to another city of British Columbia.'

The second type has the auxiliary and the beginning part of the past participle before changing the auxiliary to *être* and then pronouncing the full past participle, as in (21).

(21) *J'ai vu un camion* ***qui a p-, qui est parti*** *avec un lampadaire, avec une lumière.* (Dominic)
'I saw a truck that c-, that carried away a street light, a light.'

The third type has the speaker almost pronouncing the full phrase with *avoir,* but self-corrects to *être* before doing so, as in (22).

(22) *Apprendre à conduire euh, ça,* ***ça a arriv-, c'est arrivé*** *assez jeune.* (Sylvain)
'Learning to drive huh, it, it happen-, it happened (when I was) relatively young.'

The fourth type has intervening material appearing only between the self-correction to *être* and the past participle, as in (23).

(23) ***T'as, t'es déjà allée*** *à Sintra?* (Sabrina)
'Have you, have you ever been to Sintra?'

The fifth type has intervening material surfacing after the *avoir* auxiliary and being repeated between the self-correction to *être* and the past participle, as in (24).

(24) *Ben nous aussi* ***on a tout de suite, on est tout de suite sortis*** *de Palerme, mais mal-, malgré ça on a pris une route où c'était plein de déchets.* (Madeleine)
'Well we also immediately, we immediately left Palermo, but des-, despite that we took a road that was full of garbage.'

The sixth type has the intervening material only surfacing after the *avoir* auxiliary and not repeated between the self-correction to *être* and the past participle, as in (25).

(25) *C'était l'hiver, pis euh,* ***j'ai sûrement, je suis arrivé*** *ben trop vite, l'auto a jamais tourné.* (Charles-Antoine)
'It was during winter, and huh, I probably, I probably arrived way too fast, the car never turned.'

The seventh type has a subject pronoun change between the false start with *avoir* and the self-correction to *être*. Example 26 also includes a type of subject change, from a subject clitic (*on*) to a strong pronoun (*moi je*).

(26) *Pis là on a trippé là,* ***on a, moi je suis restée*** *comme un trois semaines, un mois, à Madagascar, dans la capitale, parce que je travaillais avec euh des jeunes de la rue.* (Rachel)
'And [discourse particle] we had a blast there, we, I stayed like for three weeks, a month, in Madagascar, in the capital, because I was working with huh street youth.'

The last type has tense change between the false start with *avoir* and the self-correction to *être*, as in (27).

(27) *Pis euh* ***on avait, on est allés*** *à Tolède, en Espagne.* (Carl)
'And huh we had, we went to Toledo, in Spain.'

3.2.1.2. Avoidance mechanisms

In the 2016 corpus, it was noticed that two different types of 'strategy', or avoidance mechanisms, were being used by speakers to convey the same meaning without actually using a periphrastic tense with an *Ê*-verb. The first strategy concerns the verb *aller*: using *avoir été* in

place of *être allé*. The second one concerns the *passé composé* tense: the use of historic present. These collected tokens were analysed separately in order to determine whether they represent yet another way of expressing past actions without resorting to the *être* auxiliary verb and/or whether speakers simply use them for pragmatic discourse functions.

In the 2016 corpus, it was also noted that more than half of the speakers (26 out of 48) used reflexes of *avoir été* in contexts where *être allé* might be expected. An example of this replacement can be found in (28).

(28) *Pis le soir, hier soir,* ***j'ai été*** *voir Maman, euh j'****ai été*** *voir Grand-Maman, à l'hôpita-, au CHSLD*[120]. (Gaëtan)
'And in the evening, I went to see Mom, uh I went to see Grandma, in the hospita-, in the CHSLD.'

Since the auxiliary that always surfaces with the verb *être* is *avoir,* I wanted to verify whether this avoidance mechanism might be dictated by the same variables as the other verbs studied here.

In colloquial/spoken French, the periphrastic past tenses of *aller* (especially the *passé composé* and *plus-que-parfait*) are often replaced with those of *être*. The *Guide du rédacteur*, an online writing aid provided by *TERMIUM Plus*, the Government of Canada's terminology and linguistic data bank, however criticizes using *avoir été* in such contexts[121] (see examples in 29):

(29a) « *Je* ***suis allée*** *travailler ce matin. (plutôt que : J'****ai été*** *travailler)* »
I am gone to work this morning. (instead of : I have been to work)
'I went to work this morning'

(29b) « *Nous* ***sommes allés*** *voir le dernier film d'Almodovar. (plutôt que : Nous* ***avons été*** *voir le dernier film d'Almodovar.)* »
We are gone to see the new Almodovar movie. (instead of : We have been to see the new Almodovar movie)
'We went to see the new Almodovar movie'

[120] *Centre d'hébergement de soins de longue durée* (CHSLD) 'long-term care facility/nursing home'.
[121] https://www.btb.termiumplus.gc.ca/tpv2guides/guides/clefsfp/indexfra.html?lang=fra&lettr=indx_catlog_a&page=9DI0SfNmDWYc.html [webpage accessed on January 24, 2020]

This feature is not just limited to Canadian French. While *Le Bon Usage* mentions that this replacement of *aller* with *être* in the past periphrastic tenses is allowed in la "langue courante" (Grevisse 2011: §803), l'Académie française, in its online writing tool platform (*Dire, Ne pas dire*), indicates the the verb *être* is "trop souvent employé à la place du verbe *Aller,* qui doit toujours lui être préféré dans le sens de *Se rendre (à), Rendre visite (à), de Convenir, ou de Se porter, Se comporter, Se conduire*".[122]

(30a) « *Il* ***a été*** *à Paris* pour *Il* ***est allé*** *à Paris* »[123]
'He went to Paris'

(30b) « *Nous* ***avions été*** *chez eux* pour *Nous* ***étions allés*** *chez eux* »
'We went to their house'

(30c) « *Ce rôle lui* ***aurait*** *bien* ***été*** au lieu de *Ce rôle lui* ***serait*** *bien* ***allé*** »
'This role would have suited him'

(30d) « *Comment ça* ***a été*** *?* au lieu de *Comment* ***est****-ce* ***allé*** *? Cela* ***est****-il bien* ***allé*** *?* »
'How did it go?'

Frei (1929: 87) indicates that grammarians used to indicate that *être allé* and *avoir été* had different meanings and this is why the verbs could not be used as semantic equivalents. It was said that *être allé* would refer to a one-way journey and *avoir été* for a return one. Frei however criticizes this approach:

> La manie des puristes et des grammairiens de chercher dans certaines fluctuations de l'usage des nuances sémantiques subtiles, relève du même besoin que la bifurcation des synonymes et n'en est que l'exagération. C'est ainsi qu'ils veulent voir une différence entre *je suis allé* (« aller simple ») et *j'ai été* (« aller-retour »), […] Personne au monde n'a jamais su où ils prenaient tout cela.[124]

[122] 'too often used instead of the verb *aller,* which should always be preferred when meaning 'to go', 'to visit', 'to suit', or 'to be (feeling)', 'to behave' '.
[123] http://www.academie-francaise.fr/jai-ete-ca-ete [webpage accessed on January 24, 2020]
[124] 'The terrible habit that purists and grammarians have of searching for certain fluctuations of use based on subtle semantic nuances stems from the same need as the bifurcation of synonyms and is an exaggeration of it. That is how they seek a difference between *je suis allé* ('one-way journey') and *j'ai été* ('return journey'), […] No one in the world has ever known where they got that idea from.'

I also wanted to check which social variables and linguistic contexts favoured the surfacing of the historic present for these intransitive *Ê*-verbs. Two examples of historic present use are given in (31), one in the recent past (31a) and one in the very distant past (31b).

(31a) [*Hier*] *j'arrive à 9 heures pile, je sais que le camp* [*de jour*] *commence à 9 heures.* (Mario)
'[Yesterday] I arrived at 9 o'clock sharp, I knew that the [day] camp started at 9 o'clock.'

(31b) *J'avais, je pense, je devais avoir cinq ans, notre gardienne était française pis elle dit qu'elle amenait des gâteaux, toi. Moi j'étais s-, convaincu que ça s'en venait la grosse affaire, pis là y'**arrive** des biscuits, j'étais comme : « Qu'est-ce c'est ça, tabarnak ? »* (David)
'I was, I think, I must have been five years old, our babysitter was French and she says she was bringing cakes, no less. I was s-, convinced that it was going to be a big deal, and then it's cookies that arrive (instead), I was like: "What the fuck is that?".'

3.2.2. Sociolinguistic independent variables

It was expected that, after transcribing the data, enough tokens per cell would be obtained to test at least the effect of the variables 'Gender', 'Age', and 'Socioprofessional status' quantitatively with the help of the statistical software *Shiny Rbrul* (Johnson 2017), which will be discussed in detail in §3.5.2.

3.2.2.1. Individual speaker

The study also attempts to determine how much intra- and interspeaker variation can be observed in the auxiliation patterns of the 48 participants, and whether individual speakers are categorical users of a variant or another, potentially depending on the verb.

3.2.2.2. Gender

According to previous studies, male speakers use the non-standard form (with *avoir*) slightly more often than do women (see §2.4.2.2.1.). In order to analyse the impact of gender on the

data, both men (24 participants) and women (24 participants) are included in the study. The tokens were coded as being uttered by either a male or a female speaker.

3.2.2.3. Age

As stated in §2.4.2.2.2., Sankoff & Thibault observed that in 1971 younger speakers tended to generalize *avoir* slightly more than their elders, which would potentially indicate a change in progress at the time. In my discussion of the sampling grid, in §3.1.2., it was mentioned that this study examines two age groups: 18-35-year-olds (24 participants) and 36-65-year-olds (24 participants). At first glance, these categories encompass a considerable age range, but these initial two groupings served to ensure that a full age range was represented when selecting participants. A complementary continuous variable is also introduced as a more fine-grained approach, not available when Sankoff & Thibault published their paper in 1977. This variable was therefore coded twice, once as age categories (young vs old) and once as continuous, with the actual ages of the speakers, with birth years ranging from 1954 to 1998.

3.2.2.4. Socioprofessional status

In order to determine the appropriate number of divisions of socioprofessional status, it was essential to make sure that they would be specific to the Montréal situation. A classic methodological issue that arises when conducting a 'first-wave'[125] sociolinguistic study is how to define the macro-category of socioprofessional status (or social class, socioeconomic status, etc.) and how to assign speakers accurately to the appropriate classification. Schilling (2013: 47) explains this problem very clearly:

[125] As Eckert states: "The studies in this tradition use survey and quantitative methods to examine the relation between linguistic variability and major demographic categories (class, age, sex class, ethnicity)." https://web.stanford.edu/~eckert/thirdwave.html [webpage accessed on June 4, 2020] (see also Eckert 2012 for a detailed account of the three different waves in sociolinguistic research)

> In Western society at least, we probably tend to think first of some mixture of economic position and social status, perhaps ascertained by measures such as income, occupation, education, and type of residence/neighborhood. However, economic worth does not always neatly correlate with social standing, and sometimes people who can be grouped into the same economic class may have different social statuses. Hence, for example, in a given community in the US it may be the case that a university professor and an experienced plumber have the same income; however, the professor will be accorded higher status because their occupation is more prestigious.

Sankoff & Thibault (1977) initially made use of a six-level geographical scale based simply on the average income of the man of the household in the area of residence of the subject (Cedergren 2018) but preferred resorting to the linguistic market index.[126] In the pilot study (Rea 2014), it proved difficult in some cases to allocate speakers with certainty to a class based solely on their type of occupation. Furthermore, in the majority of cases the level of education correlated with the type of occupation, but combining the two parameters would produce a more accurate picture when there was a mismatch between the type of occupation and the level of education completed. Based upon Ash's account (2013) of the various possible methodologies for the attribution of socioprofessional status, the following studies therefore served as models: Labov's work on Philadelphia in the 1970s (Labov 2001), which used an index combining education, occupation, and residence value; and Trudgill's (1974a) study of Norwich where he set up a social class index based on six parameters: occupation, father's occupation, income, education, locality, and housing.

Asking sensitive questions, such as yearly income or residence value, of people I knew would potentially have impeded the natural flow of the conversation by creating an intrusive environment. Rather, a hybrid index combining four parameters from both Labov (2001) and

[126] As stated in §2.4.2.2.4., the LMI is a scale developped by Sankoff & Laberge (1978) calculating how the speakers' activity within the market of linguistic exchanges is related to the speakers' use of the legitimate speech for this symbolic market. Sometimes used as a proxy for the study of SPS, it can measure the extent to which the standard language is valued in people's economic life. While Sankoff & Thibault (1977) had eight Québécois sociolinguists on hand to review the occupational history of all their participants, my calculation of the LMI would be slightly impressionistic and therefore not as precise as a socioprofessional class index that includes education history and current occupation. In many cases during the interviews, my speakers volunteered information regarding their previous occupations, but this was not done uniformly. For these reasons, the LMI is not used as an analytic tool in this study.

Trudgill (1974a) was created, where each parameter was rated with an index score from 1 to 6, and the scores for all four categories were then summed up. The four parameters comprise: type of occupation of the speaker, type of occupation of the parents/partner,[127] level of education, and locality. Each parameter consists of six subdivisions so that the total score can be divided by three (LOW, MID, HIGH) and so that all four parameters have the same weight in the total. After the four scores are summed, a total between 4 and 10 indicates membership of the lowest socioprofessional status, between 11 and 17 the middle class, and between 18 and 24 the highest socioprofessional status (see Table 3.5 below). I am confident that these social class divisions accurately reflect the socioprofessional situation in Montréal rather than being simply contextual (based on the differences between the individual informants).

Type of occupation: For their work on the Montréal French inflected future, Wagner & Sankoff (2011) employed the following 6-point occupational scale designed by Thibault & Vincent (1990) for the 1984 Montréal French corpus in order to create three levels of socioprofessional status (levels 1-2 representing LOW SPS, levels 3-4 representing MID SPS, and levels 5-6 representing HIGH SPS). Their model, presented in §2.4.2.2.3. and which features again below, is the one that I have followed in my study:

6 - Upper class: "Liberal professions" (e.g., law, medicine) and owners of businesses

5 - Upper middle class: Employed university graduates

4 - Middle class: Technicians, managers, foremen

3 - Upper working class: White-collar workers

2 - Lower working class: Blue-collar workers

1 - Lower class: No stable employment

[127] When the participant did not have a partner, the two scores of the parents' occupations were averaged.

Having the type of occupation occur twice in the index (with the inclusion of the type of occupation of the parents/partner of the speaker) allowed me to give stronger weight to this variable, as in most previous studies the type of occupation was the sole indicator of socioprofessional status.

Level of education: Since the two highest levels of occupation imply a certain level of education, it was thought best to group these two factors into the calculation of the SPS. The following scale, specific to the school-system in the province of Québec, was created:

6 - has a master's or a doctoral degree

5 - has an undergraduate degree

4 - has finished university-preparatory CÉGEP (2-year program)

3 - has finished technical CÉGEP (3-year program)

2 - has finished secondary school

1 - has not finished secondary school

Locality: As Ash (2013: 356) explains: "Trudgill's 'locality' is the neighborhood in the city of Norwich where the speaker lives, subjectively ranked for desirability on the basis of the researcher's native knowledge of the city." With this classic and influential study as a model, ten Montréalers (M1-M10) in my social circle were asked to rank the 48 speakers' neighbourhood/municipality of living in decreasing order of desirability. The ten ranking scores for each of the 17 localities were then averaged. The localities featured in the corpus was then divided by three: the first four were considered the most desirable (receiving the score of '3') and the bottom eight the least desirable (receiving the score of '1'). The middle five received a score of '2'. In order to determine where to put the cut-off between the

localities receiving the scores of '3' and '2', I decided to establish a division where a bigger gap occurred in the score (4,9 vs 7,3), i.e. between Notre-Dame-de-Grâce (NDG) and Verdun. For the second delimitation – between the localities receiving the scores of '2' and '1' – the cut-off was placed between Parc-Extension and Laval because all localities appearing below (and inclusive of) Laval are not located on the Island of Montréal but on its North and South shores, which carries a certain social stigma.[128] The results are presented in Table 3.3.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	Average rank	Desirability score
Outremont	1	1	1	2	2	1	1	1	2	1	**1,3**	**3**
Plateau Mont-Royal	2	2	2	1	1	3	2	5	1	3	**2,2**	
Ville-Marie	3	4	6	4	5	5	3	8	3	2	**4,3**	
NDG	4	3	3	7	7	2	8	2	6	7	**4,9**	
Verdun	16	7	7	3	6	6	4	13	5	6	**7,3**	**2**
Côte-des-Neiges	12	5	4	8	8	4	6	16	4	9	**7,6**	
Longueuil[129]	5	8	8	9	10	8	10	3	8	10	**7,9**	
Hochelaga-Maisonneuve	17	6	5	5	3	7	5	17	10	5	**8,0**	
Parc-Extension	13	11	10	6	4	10	7	10	16	4	**9,1**	
Laval	8	9	9	15	9	9	14	4	11	11	**9,9**	**1**
Boucherville	14	14	13	11	12	14	9	7	9	8	**11,1**	
Beloeil	6	10	14	14	16	12	12	11	7	14	**11,6**	
Brossard	10	13	11	10	11	11	13	14	13	13	**11,9**	
Chambly	7	15	12	13	14	13	11	6	12	16	**11,9**	
Repentigny	15	12	16	16	13	15	15	9	17	12	**14,0**	
Candiac	11	17	15	12	15	16	17	15	15	17	**15,0**	
Mirabel	9	16	17	17	17	17	16	12	14	15	**15,0**	

Table 3.3 Localities ranked in decreasing order of desirability

[128] Moreover, while certains parts of Laval are very wealthy, my speakers from Laval did not live in such areas.
[129] Even though Longueuil is not situated on the Island of Montréal but on the South shore, it is possible that this locality is viewed by Montréalers as more "desirable" than Laval (despite their similar proximity to the island), because Longueuil has been connected to Montréal by subway since its inauguration in 1966-1967 for the Universal Exposition of 1967, while Laval was only connected to the island by subway in 2007.

Additionally, in order to account for the fact that some speakers rent properties rather than own them, which contributes to determining someone's position in the socioprofessional scale, it made sense to construct a 6-point scale that combined the level of desirability of the neighbourhood where they lived (out of 3: '1' being the least desirable and '3' the most) multiplied by 1 if the speaker was renting the property and by 2 if he or she owned it. The results of such calculations are presented in Table 3.4.

Speaker pseudonym	Landlord/ Tenant	Landlord vs Tenant score	Locality desirability score	Total locality score
Gaëtan	L	2	3	6
Christine	L	2	3	6
Richard	L	2	3	6
Martin	L	2	3	6
Madeleine	L	2	3	6
Guylaine	L	2	2	4
Florence	L	2	2	4
Justin	L	2	2	4
Paul	L	2	2	4
Steve	L	2	2	4
Mario	L	2	2	4
Rachel	T	1	3	3
David	T	1	3	3
Jean-François	T	1	3	3
Mathieu	T	1	3	3
Marie-Jeanne	T	1	3	3
Charles-Antoine	T	1	3	3
Sabrina	T	1	3	3
Alexis	T	1	3	3
Denis	T	1	2	2
Annouck	T	1	2	2
Nathan	T	1	2	2
Dominic	T	1	2	2
Martine	T	1	2	2
Francis	T	1	2	2
Marc	L	2	1	2
Yves	L	2	1	2

Denise	L	2	1	2
Carl	L	2	1	2
Sophie	T	1	2	2
Virginie	T	1	2	2
Johanne	L	2	1	2
Annie	L	2	1	2
Maxime	T	1	2	2
Jessica	T	1	2	2
Jacynthe	L	2	1	2
Sylvain	L	2	1	2
Sara	L	2	1	2
Caroline	T	1	2	2
Amélie	T	1	2	2
Philippe	T	1	2	2
Linda	T	1	1	1
Marie-Lou	T	1	1	1
Joël	T	1	1	1
Hugo	T	1	1	1
Kim	T	1	1	1
Marie-Laurence	T	1	1	1
Julie	T	1	1	1

Table 3.4 Detailed locality score by speaker

It has proved quite challenging to find participants belonging to the lowest SPS level. It might have simply been the case that fewer people now fit the occupational criteria established by Thibault & Vincent (1990) for belonging to the lowest SPS or having only completed the lowest level of education. Since it has been obligatory from 1988 to receive formal education at least until the age of 16 (Allard 2013), it is practically impossible to find 'young' speakers today who will only have completed primary school education. Moreover, in 1971, CÉGEPS (junior colleges instituted during the Quiet Revolution[130]), had only been founded four years prior and were nowhere near as common as they are today. In their 2007 paper on change

[130] The Quiet Revolution is a period of intense socio-political and socio-cultural change in Québec mostly taking place in the 1960s. It is characterized by an effort, by the provincial government, to take more direct control over healthcare and education, previously administered by the Catholic Church, by the creation of a welfare-state, by the nationalization of electricity production, and by an intensification of the sovereigntist movement.

across the lifespan, Sankoff & Blondeau discuss the social class stratifications that had been established in 1971 as well as in 1984, and they noticed that the working-class category of 1971 included many more members of the "underclass" – "people who in 1971 were unemployed, some of whom were illiterate, had a year or less of formal schooling, and had lived their entire lives on welfare" – compared to 1984 (Sankoff & Blondeau 2007: 567). Moreover, while the twelve "younger speakers who were added to the sample in 1984 were a good match for their 1971 counterparts in age", Sankoff & Blondeau reported an increase in the mean Linguistic Market scores (the extent to which the standard language variety is valued in people's daily life). They attributed this in part to overall socioeconomic change in the community: "by 1984, the economic recession of the early 1970s had turned around, and more families could afford to keep their teenagers in school longer" (Sankoff & Blondeau 2007: 574).

In addition to the apparent rarity of LOW SPS speakers, or to the fact that their classification features have evolved since 1971, it may have been difficult for me to find LOW SPS speakers because of my own class background: I have very few personal contacts who fit such selection criteria, and it might be the case that I wrongly labelled some MID SPS speakers as LOW SPS speakers, and some HIGH SPS speakers as MID SPS ones. The detailed calculations of SPS scores for each speaker is presented in Table 3.5.

Speaker pseudonym	Occupation	Education	Parents/partner occupation	Locality	Total SPS score	SPS
Martin	6	6	5	6	23	H
Madeleine	6	5	6	6	23	H
Gaëtan	5	6	6	6	23	H
Florence	6	6	6	4	22	H
Jean-François	6	6	6	3	21	H
Charles-Antoine	6	6	6	3	21	H
Sabrina	6	6	6	3	21	H
David	5	6	6	3	20	H
Guylaine	4	5	6	4	19	H
Christine	3	5	5	6	19	H
Richard	5	5	3	6	19	H
Paul	5	6	4	4	19	H
Justin	4	5	5	4	18	H
Virginie	6	5	5	2	18	H
Marie-Jeanne	4	5	6	3	18	H
Amélie	5	5	6	2	18	H
Steve	5	5	3	4	17	M
Maxime	5	6	4	2	17	M
Jessica	4	6	5	2	17	M
Marc	4	6	4	2	16	M
Mario	4	4	4	4	16	M
Rachel	3	5	4	3	15	M
Johanne	4	5	4	2	15	M
Mathieu	4	5	3	3	15	M
Sophie	2	5	5	2	14	M
Alexis	4	3	4	3	14	M
Annouck	4	5	2	2	13	M
Annie	4	5	2	2	13	M
Marie-Lou	4	5	2	1	12	M
Denise	3	5	2	2	12	M
Nathan	4	2	3	2	11	M
Denis	2	3	4	2	11	M
Caroline	3	2	3	2	10	L
Joël	2	3	4	1	10	L
Jacynthe	2	3	3	2	10	L
Martine	3	4	1	2	10	L
Carl	2	2	4	2	10	L
Sara	3	3	2	2	10	L
Sylvain	2	2	3	2	9	L
Hugo	4	2	2	1	9	L

Francis	1	2	4	2	9	L
Dominic	1	1	5	2	9	L
Yves	2	2	3	2	9	L
Julie	3	2	3	1	9	L
Kim	3	2	2	1	8	L
Marie-Laurence	2	1	2	1	6	L
Philippe	1	1	2	2	6	L
Linda	1	1	2	1	5	L

Table 3.5 Overall calculations of SPS scores by speaker

3.2.2.5. Place of childhood

As mentioned in §3.1.1., it proved difficult to find native speakers of Montréal French and, as a consequence, it was decided that the corpus would include speakers not only from all over the *Grand Montréal*/Greater Montréal area but also native Québécois French speakers who might not have been born and raised in Montréal, but come from neighbouring regions, have lived in Montréal for most of their economic life and are now established in the *Grand Montréal,* and work or study full-time on the Island of Montréal. This sampling decision allowed me to test whether the place of childhood of the speakers has an influence on auxiliary choice. The metadata collected in the personal information questionnaire included the place of childhood of the speakers and was coded as being within the Greater Montréal area or outside of it.

3.2.2.6. Language and dialect contact

Most studies on auxiliary alternation in French have been carried out in North America and in primarily English-speaking communities (in Ontario, in the Western and Maritime Provinces of Canada, and in New England and Louisiana, United States. For the full list, see introduction to §2.4.2.) where French does not have an official status as opposed to Québec.

While some studies have shown that this alternation exists in a few French-speaking regions of Europe, it would be relevant to analyse the impact of this variable in Montréal, even though French enjoys an officially accepted status there, since spoken Québécois French is so readily associated with an important use of Anglicisms (mostly semantic, lexical, and syntactic ones) and since HAVE is the sole auxiliary verb of the present perfect tense in English. There is no doubt that Québécois French has been and still is influenced by the close contact with the English language.

According to Poplack (1997), one of the ways to reveal the effect of contact with English is to observe the behaviour of speakers who have the best knowledge of English. I therefore analysed the impact of the self-reported use of English of the 48 speakers: a question about the frequency of usage of English (*Do you speak English at work and/or at home on a daily basis?*) was included in the personal information questionnaire (see Appendix 3B). It is therefore assumed that the speakers who use English on a daily basis have the highest proficiency. The study also tests for the different types of bilingualism (native/simultaneous vs acquired/consecutive). This last element was not part of the personal information questionnaire, but in most cases information about if, when, and how the informants had learned to speak English was obtained during the course of the sociolinguistic interviews. In a few rare cases where this information was not made available during the interview, the participants were contacted afterwards in order to establish when and how they had learned to speak English. The tokens were therefore coded first based on whether the speaker was an English bilingual or not, and then a second time based on whether that bilingualism was simultaneous (native bilingual) or sequential (acquired bilingualism).

A question about exposure to other languages is included in the personal information questionnaire in order to test whether contact with a language that only makes use of a single auxiliary, for example a Romance language such as Spanish which only uses auxiliary HAVE,

might influence *avoir* selection. A question about exposure to other French dialects is also included in the personal information questionnaire in order to test whether patterns in my data replicate the results of Stelling (2011: 11) lending "support to the notion that contact with varieties of French other than the source dialect" discourages use of *avoir*. Contact with languages other than English as well as contact with other French dialects have not been included in the multivariate analysis, but rather analysed with descriptive statistics (see §4.2.5.1.2. and §4.2.5.1.3.) because the data collected through the questionnaire were not consistent across all participants, since it transpired that their interpretation of what counted as a contact varied considerably.

3.2.2.7. Stylistic variation

In order to test for the influence of (interspeaker) stylistic variation, the study analyses two parameters to create an index scale of style. The speakers were first divided into two groups based on whether or not they had been interviewed in a pair. They received the score of '1' if they had been interviewed individually and '2' if they had been interviewed in a pair. The reason behind this classification is that, based on work by Labov (1972), one can assume that speakers pay less attention to their speech when they engage with their peers, as mentioned earlier in §3.1.3. Subsequently, an scale of familiarity with three levels was created, in which three possible types of relationship were envisaged. The informant either:

1 - Knows the interviewer through a distant personal contact of the interviewer

2 - Knows the interviewer through a close personal contact of both the speaker and the interviewer

3 - Knows the interviewer personally

It might be the case that, when taking into account other sociolinguistic variables, speakers who know the interviewer personally (or through a close personal contact) produce more *avoir* tokens than the speakers belonging to the first category. In the same line of reasoning, it is possible that speakers who have been interviewed in a pair, and therefore used more informal speech, produce more *avoir* tokens. Their initial pair score of '1' or '2' was then multiplied by 1 if the participant had a distant personal connection to the interviewer, by 2 if the participant knew the interviewer through a close contact, or by 3 if the participant knew the interviewer on a personal level. This calculation produced a scale with five possibilities (1, 2, 3, 4 or 6): from a score ranging from 1 (where it is likely that participants were the most careful with their speech, say, if a participant who did not know the interviewer beforehand was interviewed on her own) to 6 (where it is likely that participants were the least careful with their speech, for example if two close contacts of the interviewer were interviewed together). The *avoir* rates of the speakers in each of the five levels of formality were then averaged.

Stylistic variation is not included in the multivariate analysis, but rather analysed with descriptive statistics because the calculations of this factor are not as robust as the other factors, and it was felt that data that is more approximative could skew the statistical model results. These matters are further discussed in §4.2.5.1.1.

3.2.3. Linguistic independent variables

3.2.3.1. Lexical item

As shown in §2.4.2.3.1, *Ê*-verbs always select *avoir* in completely different proportions, and I compared the *avoir*-selection rates of the various intransitive verbs to Sankoff & Thibault's rankings in order to see whether the lexical effect that they had observed still holds valid. Sankoff & Thibault make no mention of whether they included iterative forms of the verbs

in their total calculations, but they only seem to treat *revenir* as a separate lexical item from *venir* (Sankoff & Thibault 1977: 96). For this reason, my tokens were coded twice: once with iterative and non-iterative versions coded separately, and once with them coded together. The complete list of verbs and iterative forms studied here is presented in Table 3.6.

Non-iterative *Ê*-verbs	**Iterative *Ê*-verbs (where relevant)**
aller	
arriver	
déménager	*redéménager*
demeurer	
descendre	*redescendre*
devenir	*redevenir*
(r)entrer[131]	*re-rentrer*
monter	*remonter*
partir	*repartir*
passer	*repasser*
rester	
sortir	*ressortir*
tomber	*retomber*
retourner	
venir	*revenir*

Table 3.6 Complete list of intransitive *Ê*-verbs and their iterative forms

3.2.3.2. Verb meanings

Out of all the intransitive verbs studied, some can have both a 'core' meaning, usually motional or relating to a change of state, and a figurative meaning, as explained in §2.4.2.3.2. Examples of figurative and/or lexicalized meanings given by Sankoff, Thibault, & Wagner (2004) include *sortir avec quelqu'un* 'to date/to go out with someone' and *venir au monde* 'to be born' (lit. 'to come into the world'). The study also seeks to determine whether the 'core' meaning of the verbs *sortir*, *tomber, passer*, and *rentrer* (usually motional or relating to a change

[131] While *rentrer* is iterative in its morphology, it very rarely carries an iterative meaning in Montréal French. More detail about this is provided in §3.4.

of state) selects *avoir* in the same proportions as their figurative meanings. An example, taken from the 2016 corpus, of the figurative use of *sortir* can be found in (32).

(32) *Un que* ***j'ai sorti*** *avec, de l'âge, depuis l'âge de, de 15 à 18 ans.* (Sara)
'One that I've gone out with, from the age, since the age of, of 15 to 18 years old'

The verb *tomber* can have multiple figurative uses. Some are outlined in (33a-c). They can comprise 'to become suddenly', 'to get (pregnant)', 'to work out well/badly', 'to come at the right/wrong time, to be (un)lucky', 'to fall (in love)', etc.[132]

(33a) *Pis euh dans le fond la maison* ***a tombé*** *à vendre, parce que la madame pouvait pu s'occuper de la maison.* (Carl)
'And huh basically the house came on sale, because the lady couldn't take care of the house anymore'

(33b) *Écoute, moi là, c'est fou, parce que quand même j'ai, je me, je me, tsé je faisais pas attention disons, ok : j'****ai*** *jamais* ***tombé*** *enceinte.* (Julie)
'Listen, me, it's crazy, because even though I have, I, I, you know I wasn't being careful let's say, ok: I never got pregnant.'

(33c) *Le ménage je le fais pas aujourd'hui ! Ben regarde ça* ***a-****tu ben* ***tombé !*** (Linda)
'The cleaning I'm not doing it today! So look how well that worked out!'

An example of a figurative use of *passer* is the idiomatic expression *passer proche (de mourir)* 'to come close to dying', as shown in (34), compared with other uses.

(34) *Avant ça oui, la fièvre typhoïde, ça c'est, c'est une maladie, mais euh c'est vrai j'****ai passé proche*** [*de mourir*]. (Philippe)
'Before that yes, typhoid fever, that's a, that's a disease, but huh it's true that I came close [to dying]'

[132] Some of these collocations are more frequent than others.

The study also looks at the specific use of the verb *rentrer* in the context of bicycle and car crashes, as illustrated in example 35, as opposed to other uses.

(35) *Y'aurait eu un char, y'**aurait** peut-être **rentré** dedans pareil, mais le gars aurait été protégé dans son char, tsé y'aurait pas revolé.* (Steve)
'If there had been a car, he would have maybe hit it anyways, but the guy would have been protected inside his car, you know, he wouldn't have gone flying'

The various verb meanings have not been not integrated in the multivariate analysis, but rather analysed with descriptive statistics (see §4.2.5.2.1.) because the statistical model could not converge when that many additional factors were included.

3.2.3.3. Frequency of use

The study analyses the relative frequency of each verb, i.e. the total number of periphrastic occurrences in the corpus (rather than occurrences per X number of words), and their iterative forms as they appear in the corpus, in order to test whether their rates of *avoir* tokens correlate with their respective frequency of use on a linear graph.[133] It also tests whether high-frequency verbs continue to disfavour *avoir* usage, as was the case in the 1971 corpus.

Frequency of use is not included in the multivariate analysis, but rather analysed with descriptive statistics (see §4.2.5.2.2.), because the overall frequency of each verb, including all synthetic tenses, has not been calculated since it is only the compound tense tokens and historic present tokens that have been transcribed. It was therefore thought best to exclude such incomplete data from the statistical model so as not to skew its results.

[133] Future work on auxiliary alternation could make use of a frequency dictionary, for example the Québécois French frequency index dictionary of Beauchemin, Martel, & Théoret (1992), rather than using the relative corpus frequency, since it is possible that my corpus verb frequencies are not an accurate reflection of the verbs in 'real' use.

3.2.3.4. Possibility of parallel adjectival use

A survey of previous studies shows that the literature was inconsistent with regard to this usage, as mentioned in §2.4.2.3.6., because the influence of the possibility of parallel adjectival use has been tested in three different ways. A summary of these findings for each verb is presented in Table 3.7, together with my own grammaticality judgements as a native speaker of Montréal French.

I accepted the verbs that could be used as parallel participial adjectives with copula and adverb *maintenant* – the method used by Canale *et. al.* (1978) – because the resulting sentences sounded much closer to informal spoken French than the ones resulting from the method used by Sankoff & Thibault (1977, 1980) (adjectival use possible without copula, which sounded very formal to me). Moreover, I disagreed with how Sankoff & Thibault had classified the verbs *arriver* and *revenir* which, in my judgement, allowed parallel adjectival use even without copula.

A dash '-' indicates that the verb has not been examined in any study, an 'X' that no alternation was recorded in the study, a '#' that too few tokens were collected to be studied, an 'N' that the past participle cannot be used as an adjective according to the author, and lastly a 'Y' that it can. The '?' indicates the only case, *rester*, where, in my judgement, there is a mismatch of test results in the possibility of adjectival use of the verb whether this variable is tested with the method of Canale *et al.* (1978) or would have been tested with the method of Sankoff & Thibault (1977), both illustrated above in §2.4.2.3.6.[134]

[134] For me, the sentence *L'enfant restée à la maison était malade* 'The child who stayed home was sick' is grammatical but **Elle est restée maintenant* 'She is stayed now' is not. Verb *rester* is therefore excluded from the analysis of that variable, presented in §4.2.5.2.3.

Verb	Sankoff & Thibault (1977)	Canale *et al.* (1978)	Willis (2000)	Rea (2020)
Aller	N	N	-	N
Arriver	N	Y	N	Y
Déménager	Y	-	Y	Y
Demeurer	-	-	Y	#
(Re)descendre	Y	-	N	Y
Devenir	-	-	N	X
Entrer	N	-	N	#
Monter	Y	-	Y	Y
(Re)partir	Y	Y	Y	Y
Passer	Y	-	Y	Y
Rentrer	Y	Y	Y	Y
Rester	-	N	Y	?
Retourner	Y	-	Y	Y
Revenir	N	Y	Y	Y
Sortir	Y	Y	Y	Y
Tomber	Y	Y	Y	Y
Venir	N	N	N	X

Table 3.7 Verbs which allow parallel adjectival use (according to various studies on auxiliary alternation)

After adding my own grammaticality judgements as a native speaker of Montréal French, Table 3.7 shows that there is only a consensus across the studies for the verbs *aller, déménager, monter, (re)partir, passer, rentrer, retourner, sortir, tomber,* and *venir*. Table 3.7 also shows that the verbs *aller* and *venir* are the only two cases in this list where all authors agree that they do not allow adjectival use. This variable is not tested with pronominal verb tokens.

The possibility of parallel adjectival use is not included in the multivariate analysis, but rather analysed with descriptive statistics (see §4.2.5.2.3.) because there was no consensus in the field regarding how to code for this factor. It was therefore thought best to treat this factor separately so as not to include potentially incorrect data in the statistical model.

3.2.3.5. Possibility of parallel transitive use

Here verbs were coded based on whether or not they allow (direct or indirect) transitive use.[135] As stated in §2.4.2.3.4., the verbs which have transitive equivalents are *(re)partir* (Québécism), *(re)monter*, *(re)descendre, (r)entrer*, *(res)sortir*, *(re)passer*, *(re)déménager,* and *retourner*. The verbs which do not allow parallel transitive use are *aller, arriver, rester, (re)tomber* (in Laurentian French), and *(re)venir*.

When the token is a pronominal verb, the coding is based on whether the verb without the reflexive pronoun can have a transitive use. For example, tokens of *se laver* are coded as permitting transitive use because of the existence of transitive verb *laver,* but tokens of *s'enfuir* are coded as not permitting transitive use because transitive *enfuir* does not exist/is not grammatical.

3.2.3.6. Possibility of parallel pronominal use

Some intransitive *Ê*-verbs, like *sortir*, can be used pronominally, as in *se sortir* 'to get oneself out (of)', whereas some verbs, such as *arriver, rester* or *tomber,* cannot. With the case of *partir*, it is one of the intransitive *Ê*-verbs that do allow pronominal use, but it can only do so in two idiomatic expressions exclusively found in Laurentian French: *se partir à son compte* 'to set up one's own business, to become self-employed' and *se partir en affaires* 'to start a business'.[136]

The verbs were coded in a tripartite way, based on whether or not they can allow pronominal use, and whether that use is restricted to contexts where the verb is preceded by

[135] Future studies on auxiliary alternation could employ a transitivity score for each verb, rather than code the data using a binary approach (transitive use allowed vs not allowed). The score could be created by surveying the 2016 corpus to determine the level of transitivity of these verbs. This fine-grained transitivity scale could properly reflect in what proportion the verb is used transitively in a given corpus and could potentially show the differential impact of transitivity on auxiliary choice.

[136] The dictionary *Usito,* the first online dictionary of standard Québec French, mentions that the use of *partir à son compte* or *se partir à son compte* (from the English *to start on his own account*) is criticized as a non-standard synonym of *s'établir*, *s'installer*, *se lancer à son compte* 'to set up one's own business, to become self-employed', and that the use of *partir en affaires* or *se partir en affaires* (from the English *to start a firm, in a business*) is also criticized as a non-standard synonym of *créer (une entreprise, un commerce)*, *fonder (une entreprise, un commerce)*, *ouvrir (une entreprise, un commerce)*, *se lancer en affaires* "to start a business'. Both expressions are marked as Québécisms. https://usito.usherbrooke.ca/d%C3%A9finitions/partir_2#e54c (webpage visited on February 14, 2020)

locative pronoun *en*, such as *aller* < *s'en aller* 'to go away' and *(re)venir* < *s'en (re)venir* 'to come (back)' or 'to arrive'. For obvious reasons, this variable is not tested with pronominal verb tokens.

3.2.3.7. Subject person and number

The tokens are coded based on the person and number of the subject, as outlined below:

1SG *je* 'I'

2SG *tu* 'you.SG.informal'

3SG *il, elle,* NPs, etc. 'he', 'she', 'it'

1PL *on* lit. 'one', but usually referencing plural 'we.informal' and *nous* 'we.formal'

2PL *vous* 'you.PL'[137]

3PL *ils, elles,* NPs, etc. 'they.Masc', 'they.Fem'

With pronominal verb data, it was necessary to exclude all 3PL subject tokens because without the use of a phonetics speech analysis software like Praat it was very difficult to distinguish between *avoir* and *être* tokens, namely *s'ont* vs *(se) sont.*

3.2.3.8. Form of auxiliary

It was necessary to distinguish this variable from the 'Subject person and number' because the pronoun *on,* which is very rarely used in spoken Laurentian French with an indefinite singular reference, has plural reference (1PL) but a 3SG auxiliary form. Moreover, there were a few cases of mismatch in the number agreement of the subject and the auxiliary verb. In some cases, the agreement is made *ad sensum* rather than syntactically, as illustrated in (36).

[137] Pronoun *vous* can also refer to 'you.SG.formal' but no tokens of *vouvoiement* were collected with the verbs under study during the sociolinguistic interviews.

(36) *Pis j'ai de la* ***famille*** (3SG) ***qui sont allés*** (3PL) *aussi.* (Annouck)
'And I have some family that went as well.'

And in others, the mismatch appears when the subject of the verb is relative *qui,* as in (37).

(37a) *Pis dans le fond, c'est même pas* ***moi*** (1SG) ***qui est allée*** (3SG) *le porter à la garderie, c'est son père.* (Julie)
'And actually, it's not even me who took him to daycare, it's his father.'

(37b) *Moi j'ai finalisé un peu plus tard ma job, parce que j'ai eu plein d'aff-, j'ai eu plein,* ***quelques courriels*** (3PL) ***qui est rentré*** (3SG) *en même temps.* (Mathieu)
'I finished work a bit later, because I had many thin-, I had many, a few emails that arrived all at the same time.'

With pronominal verb data, it was again necessary to exclude all 3PL verb tokens for the same reason mentioned above.

3.2.3.9. Tense and mood

As stated in §2.4.2.3.9., it has been shown that verb tenses influence auxiliary data in diverging ways: for Sankoff (2019: 207) tenses other than the *passé composé* favoured *avoir,* while for Roussel (2016) and Renaud & Villeneuve (2008) more *avoir* use surfaced with the *passé composé.* The 2016 Montréal data is coded according to the six periphrastic tenses below:

Passé composé 'past perfect'

Plus-que-parfait 'pluperfect'

Futur antérieur 'future perfect'

Conditionnel passé 'past conditional'

Subjonctif passé 'past subjunctive'

Infinitif passé 'past infinitive'

3.2.3.10. Type of subject

The study tests whether different subject types influence auxiliary selection. The tokens were coded according to the following types of subject:

ça 'it'

Impersonal *il*

Common noun NP (*ma sœur, les examens, les autres*)

Proper noun (*Madame Auger, Pierre, Maman/Papa*)

Clitic/weak pronoun [cannot be doubled] (e.g. *je, tu, il, on,* etc.)

Strong pronoun in dislocation (*moi*, *toi*, *lui, elle, eux, elles, eux autres, nous autres,* etc.)

qu'est-ce qui 'who', 'what'

quelqu'un 'someone'

Relative *qui* (including *ce qui*…) 'who', 'whom'

tout 'everything'

No subject uttered (i.e. deleted)

Left dislocations have been coded as having a clitic/weak pronoun for a subject, as in (38).

(38) ***Les roues y'****ont comme rentré, faque lui était pas capable d'ouvrir la porte.* (Rachel)
'The wheels they sort of bent inwards, so him he wasn't able to open the door.'

With the past infinitive tokens, the type of subject is coded as the one accompanying the modal verb or the adjective.

3.2.3.11. (In)animacy of the subject

As outlined in §2.4.2.3.11., Thibault & Sankoff tested for this variable in the 1997 paper (unpublished) that looked at the 1971 and 1984 Montréal data, but the results were only statistically significant for the verb *partir*, where inanimate subjects favoured *avoir*. This study

aims to test whether the subject's animacy influences auxiliary selection with all intransitive *Ê*-verbs. The tokens were coded according to whether the subject is animate or not. An example of a token with an inanimate subject is presented in (39).

(39) *Bon, quand que j'étais sur la rue ici, là, les,* ***les instructions*** *y'ont parti, au vent.* (Maxime)
'So, when I was on the street here, then, the, the instructions got caught in the wind'

Tokens with a collective noun were coded as having an animate subject if semantically it referred to humans,[138] as in (40).

(40a) *Vers 4h euh, la majorité de* ***l'équipe de vente*** *est partie, nous, nous il restait les coordonnateurs.* (Mathieu)
'At 4pm huh, the majority of the sales team left, we, we, the coordinators we stayed behind.'

(40b) *Y'a beaucoup de* ***communautés culturelles*** *qui sont allées là.* (Charles-Antoine)
'There are a lot of cultural communities that went there.'

3.2.3.12. Iterativity and morphologically-derived forms

The present study tests for the influence of iterativity on auxiliary choice, and the tokens were coded according to whether the verb appears with an iterative prefix or not. Initially, it was hypothesized that various types of morphological derivations might surface, not simply iterative forms (see §2.4.2.3.7.), however only one non-iterative morphologically-derived token was recorded: the Anglicism *passer-out,* found in (41). It was excluded from the analysis.

(41) *Je sais pas sur quoi elle* ***est passée-out*** *là.* (Mario)
'I don't know from what (substance) she passed out'

This variable is not tested for the pronominal data.

[138] Animals would theoretically also have been included in this category, but no collective nouns referring to animals have been collected.

3.2.3.13. Presence of intervening material between the auxiliary and the past participle

Since Sankoff (2019: 207) and Willis (2000: 65) showed that proximity of the auxiliary and the past participle disfavoured *avoir* use in their data, the variable 'Presence of intervening material between the auxiliary and the past participle' is tested with the Montréal 2016 data.

An example of an adverb acting as an intervening element in my corpus is presented in (42), where the auxiliary verb and the past participle are separated by adverb *tout de suite*.

(42) *J'ai* ***tout de suite*** *tombé sur un, un pâtissier français mais qui est très jeune.* (Caroline)
'I immediately stumbled on a, a French pastry chef but who is very young.'

Adverbs can also be juxtaposed and combined, as in (43).

(43) *Ouais, c'est vrai, je pense que les deux on a* ***quand même beaucoup*** *déménagé là.* (Marie-Jeanne)
'Yeah, that's right, I think that we both moved houses quite a lot.'

Certain discourse markers, particularly adverbial downtoners, can also be found between the auxiliary and the past participle, as in example 44. The most frequent are *comme* and *genre* (both meaning 'sort of, kind of, like'); *genre* is most associated with the speech of younger generations in Québécois French.

(44) *Je pense que j'avais* ***comme*** *sorti avec dans le temps parce que toutes les filles avaient des chums qui allaient à l'école avec.* (Jacynthe)
'I think that I had sort of gone out with him at the time because all the girls had boyfriends who went to school with him.'

The tokens were coded as either having intervening material, regardless of the quantity of elements separating the auxiliary from the past participle, or not.[139] Negative sentences where

[139] Future studies on auxiliary alternation could verify whether the amount of 'material' intervening between the auxiliary and the past participle plays a role in the choice of auxiliary. For example, Poplack (1992: 247), in her study of the variability of the French subjunctive, found that distance between the matrix and the embedded

there is an intervening element, i.e. negative adverb *pas,* between the auxiliary and the past participle were also coded as having intervening material.

3.2.3.14. Polarity of the sentence

While this variable was tested by Willis (2000: 40) on Ottawa-Hull French and did not prove statistically significant, I take into consideration the polarity of sentences in order to be able to distinguish the variable's potential effect from the effect of the presence of intervening material between the auxiliary and the past participle.[140] Moreover, work by Emirkanian & Sankoff (1985: 194-5) shows that polarity can influence morphosyntactic variables, such as the use of the periphrastic vs inflected future in Montréal French : negative polarity was found to be a virtually categorical environment for the use of inflected forms.[141] My tokens were coded either as being in a positive sentence or in a negative one.

3.2.3.15. Hodiernal vs pre-hodiernal action

Based on the fact that the hodiernal/pre-hodiernal distinction is a parameter that has affected tense choice in the history of French, at least in the written language, i.e. the use of the *passé composé* was once restricted to the description of actions that had taken place within 24 hours of the moment of speech (*la règle des 24 heures* 'the 24-hour rule') and the *passé simple* was to be used for all pre-hodiernal actions (Estienne 1569; Fournier 2004),[142] I hypothesized that this distinction had the potential to play a role in auxiliary choice.

---

verb disfavoured the choice of the subjunctive mood in embedded noun clauses governed by *falloir* 'to have to': when the intervening material was a single word it disfavoured the subjunctive less than when the intervening elements were parenthetical material.

[140] During the time when I was reviewing the variables to include in the study, I caught myself saying the following sentence to a friend: *Hier je* ***suis tombée****, mais aujourd'hui j'****ai pas tombé*** 'Yesterday I fell, but today I did not fall', which further convinced me to add polarity as a variable worth testing.

[141] All subsequent studies on the future in Canadian French (e.g. Deshaies & LaForge 1981; King & Nadasdi 2003; Poplack & Dion 2004; Poplack & Turpin 1999; Sankoff & Wagner 2006) have found similar results.

[142] In the Oltenian dialects of Romanian, this distinction still affects the selection of verb-forms but the patterns are (for unknown reasons) reversed in that the use of the preterite is reserved for events which have occurred in the speaker's 'today', and the periphrastic form for events that are pre-hodiernal (Maiden 2015: 48; Bertinetto & Squartini 2016: 942).

Willis (2000: 44) examined whether *avoir* was the auxiliary of choice in Ottawa-Hull French for events occurring in the distant past, which she defined as more than 24 hours ago. Willis (2000: 43) had planned to test this variable by referring to it as "proximity in the past", but excluded it from the variable rule analysis because she had too few hodiernal tokens (Willis 2000: 66). She predicted that *être* would be the auxiliary of choice for the recent past and by default *avoir* would express the distant past, based on the fact that grammarian Antonini (1753: 317) stated that auxiliary choice with certain verbs such as *aller, demeurer, monter, passer*, and *sortir* "expressed different circumstances of time, either less distant or more distant in the past": *il* ***est*** *sorti* 'he is out' is provided by Antonini as an example of someone having gone out and not yet returned, i.e. an example of proximity in the past (Willis 2000: 43).

This variable is therefore tested for the first time on Laurentian French data. The tokens in my data were coded as either being hodiernal or pre-hodiernal. An example of pre-hodiernal action is given in (45).

(45) *Mais, mais mon principal set, c'est des gros couteaux, je l'ai depuis que, depuis que j'****ai rentré*** *à l'Université de Montréal, donc à peu près* ***y'a huit ans***. (Jean-François)
'But, but my principal set, they're big knifes, I've had it since, since I started at Université de Montréal, so about 8 years ago.'

Coding the hodiernal data was very straightforward because each participant had been asked to report in detail what they had done the day prior to the interview, in the last 24 hours (see §3.1.3.): a lot of hodiernal auxiliary tokens appeared alongside each other and were easy to detect, even though they were not always accompanied by a hodiernal discourse marker such as "today" or "yesterday". While answering that question, a few participants realized along the way that they were actually reporting what they had done the day of the interview or the day prior to the day before the interview, i.e. an action that had taken place approximately within 48 hours. Fourteen such tokens have been collected.

3.2.3.16. Clause structure (main vs embedded clause)

This variable is also tested for the first time on Laurentian French data. It has been shown that language change, especially relating to word order, usually originates in main clauses, and linguistic innovations subsequently tend to spread to embedded clauses in later stages of linguistic development (Lightfoot 1991: ch. 3; Harris & Campbell 1995: 27). Subordinate clauses are therefore thought to exhibit more conservative behaviour, and it will be therefore interesting to note which auxiliary I observe more frequently in these contexts. The tokens were coded as appearing either in a main clause or in an embedded clause.

A difficulty in coding for this variable is the frequent deletion of complementizer *que,* which is a distinctive feature of Laurentian French (Sankoff *et al.* 1971; Sankoff 1974, 1980a,b; Connors 1975; Martineau 1985, 1988; Warren 1994; Roberge & Rosen 1999; Dion 2003; Nagy & Blondeau 2005), and makes it challenging to determine whether a clause is subordinated or not. This means that I have collected a lot of tokens starting with *c'est-à-dire* [*que*]*, je veux dire* [*que*]*, y'a des fois* [*que*]*, c'est juste* [*que*]*, c'est vrai* [*que*]*, je pense* [*que*]*,* and *c'est sûr* [*que*]*,* etc., where the complementizer *que* which introduces the subordinated sentence is not uttered. Such tokens, exemplified in (46), were all coded as embedded clauses.

(46a) *J'ai vraiment aimé mieux, faque c'est sûr* **[*que*]** *je me suis, je me suis inscrit à Laval.*
(Caroline)
'I really enjoyed it better, so of course I, I signed up in Laval.'

(46b) *Mais je pense* **[*que*]** *mon dernier voyage, j'ai, je suis tombée sur une toilette turque.*
(Madeleine)
'But I think [that] on my last trip, I, I came across a turkish toilet.'

(46c) *J'ai fait : « Oh my god, ça fait six mois* **[*que*]** *je me suis pas assis de même, je capote ! »*
(Sara)
'I went : "Oh my god, it's been six months since I was last able to sit like this, I'm freaking out!""

Tokens where relative pronoun *que* had been omitted were also coded as embedded, as exemplified by (47).

(47a) *Y'a une autre place* **[*que*]** *on est allés visiter euh, à un moment donné, c'est l'Île de...* (Steve)
'There is another place [that] we went to visit huh, at some point, it's the Island of…'

(47b) *J'admets que comme la plupart des cyclistes là, y'a ben des fois que j'ai pas fait... là, les stops on les fait pas, y'a des fois* **[*que*]** *j'ai passé.* (Gaëtan)
'I admit that like most cyclists [discourse particle], there are many times when I didn't do… [discourse particle], we don't stop at stop signs, there are times when I passed.'

(47c) *Seule chose* **[*que*]** *je me suis cassée dans ma vie c'est un bout de doigt.* (Denis)
'The only thing [that] I broke in my life is a bit of my finger.'

(47d) *Tu donnes, pis tu t'en sacres. Pis c'est ça* **[*que*]** *je me suis dit.* (David)
'You give, and you don't give a sh*t. And that's [what] I told myself.'

3.2.3.17. Presence of a clitic pronoun between subject and auxiliary verb

This variable is also tested for the first time on Laurentian French data. I decided to add it to the analysis after carefully examining the collected *avoir* tokens and noticing some patterns emerge. With intransitive tokens, clitic pronouns can appear between the subject and the auxiliary. With pronominal verbs, an additional clitic pronoun is added between the reflexive pronoun and the auxiliary, and it should be noted that this variable does not refer to whether there is a reflexive clitic in front of the auxiliary, as all pronominal verbs do.

With intransitive verbs, this can be illustrated in sentences such as (48), where the indirect object clitic pronoun *me* appears between the subject and the auxiliary:

(48) *Le vélo a glissé sur le côté, le guidon* ***m'a rentré*** *dans la cuisse.* (Martine)
'The bicycle slid on the side, the handlebars hit me hard in the thigh.'

Other possible pre-verbal clitics include locatives *en* and *y*, as in (49).

(49) *Mon* [*examen de*] *théorie, ben pro-, je l'ai manqué à une question, faque j'**y ai retourné**, pis après je l'ai passé.* (Linda)
'My theory [exam], well pro-, I failed it because of one question, so I went back, and after that I passed it.'

With pronominals, the clitic pronouns added to pronominal tokens can either be direct objects, such as *les*, illustrated in (50):

(50) *Moi, je **me les*** [*mes ovaires*] ***ai fait** enlever avant l'hystérectomie.* (Denise)
'I had them [my ovaries] removed before the hysterectomy.'

Or they can be combined with indirect object and partitive *en*, as in (51).

(51a) *Je **m'en aurais** jamais **douté**.* (Julie)
'I would never have guessed it.'

(51b) *Tsé j'aurais pas v-, tsé je **m'en aurais passé**, mais totalement là.* (Julie)
'You know I would not have v-, you know I could have done without it, but like totally.'

(51c) *Entoucas, j'ai t-, je prenais la rue Cousin, mais ce coup-là je prends la rue avant, c'est juste je **m'en ai** pas **rendu compte**.* (Gaëtan)
'In any case, I t-, I used to take Cousin Street, but that time I decide to take the street before it, it's just that I didn't realize it.'

3.3. Circumscribing the variable context and transcription protocols

Constructions where *être* + participle acted as a copula + adjective (illustrated in 52a), rather than as an auxiliary + past participle (52b), were not included in the study because in these forms the two variants, *être* and *avoir*, are not interchangeable (52c), as opposed to (52b):

(52a) *Il* ***est sorti*** *présentement/maintenant/en ce moment/depuis deux heures.*
'He is out presently/now/at the moment/has been out for the last two hours.'

(52b) *Il* ***est/a sorti*** *à cinq heures.*
'He went out at 5 o'clock.'

(52c) **Il* ***a sorti*** *présentement/maintenant/en ce moment/depuis deux heures.*
'He is out presently/now/at the moment/has been out for the last two hours.'

In order to justify how they circumscribed the variable, Sankoff & Thibault (1977) used, among others, temporal and non-temporal indices, indicators of duration, tenses other than the *passé composé*, etc. In the Montréal French 2016 data, the tokens where *être* is used as a copula, exemplified in (53), are therefore excluded from the analysis. Examples 53(a) and 53(b) contain temporal indicators referring to present time (*encore* 'still' and *là* 'currently, (right) now') (see Sankoff & Thibault 1977: 84, 87). Example 53(c) contains the copula in the imperfect with an indication that the action was still on-going during the event described in the main clause (see Sankoff & Thibault 1977: 87-88). Example 53(d) is also excluded from the corpus because *la première sortie* 'the first one out' is a DP.[143]

(53a) *Pis on n'****est*** *pas encore* ***sortis*** *des boîtes, deux sem-, deux mois plus tard.*
(Charles-Antoine)
'And we are still living out of boxes, two wee-, two months later.'

(53b) *Mais là elle* ***est revenue*** *à Tokyo, pis elle va rester là pour le reste de son année genre.*
(Caroline)
'But she is currently back in Tokyo, and she will stay there for like the rest of her year.'

(53c) *C'est ça, moi, je l'avais pas connu avant parce que j'****étais*** ***partie*** *à Red Deer en Alberta, un mois, deux mois.* (Annouck)
'Exactly, me, I had not met him before because I was away in Red Deer in Alberta, for a month, two months.'

[143] See Sankoff & Thibault (1977: 84-95) for additional details on how they circumscribed the variable. As explained in §2.3., this exhaustive work was conducted in the midst of the debate regarding the nature of (morpho)syntactic variables as sociolinguistic variables.

(53d) *Si genre je, j'entends une alarme, tu peux être sûre [que] je* ***suis*** *la première* ***sortie****, je veux dire y'a même pas comme…, je vais pas aider personne là.* (Florence)
'If like I, I hear an alarm, you can be sure [that] I am the first one out, I mean there is not even like…, I won't help anyone [discourse particle].'

Interestingly, Sankoff & Thibault (1980: 329) did however find in the 1971 corpus two cases where *avoir* is used in constructions analysed as "resulting state", presented in (54).

(54a) *Mettons que j'****aurais parti*** *cinq ans.* (Speaker 52, line 318)
'Let's say I'd have been away for five years.'

(54b) *Mais maintenant, eu, ... bah! ça* ***a passé****.* (Speaker 113, line 281)
'But now, uh, ... well, it's gone.'

Even though such evidence seems to indicate that in a few contexts auxiliary *avoir* can behave like a copula, Sankoff & Thibault (1980: 329) believed at the time that it would be "somewhat hasty to conclude" that *avoir* can fill the copula role. In my data, I also found at least one token, presented in (55), where *avoir* might be behaving like a copula verb, though the use of the subjunctive muddies the waters. This token was excluded from the analysis.

(55) *Faque j'ai, je comprends que tu veuilles que j'accouche pis que je…, dans ma tête, que le bébé* ***aille sorti****, mais ça me dérange pas.* (Sara)
'So I have, I get that you want me to give birth and that I…, in my head, (you want) the baby to be out now/already, but it doesn't bother me.'

The segmentations and transcriptions were done with the linguistic annotation software *ELAN*, which allowed me to separate the interview file into utterance units (segmentation process) and easily transcribe each segment in the audio file. Figure 3.1 shows a screen capture of the segmentation mode in *ELAN*, which allows the user to divide an audio file into distinct sound tracks. These 'mini' sound tracks are then made available in *ELAN*'s transcription mode (presented in Figure 3.2) which plays them back to back in order to facilitate transcription.

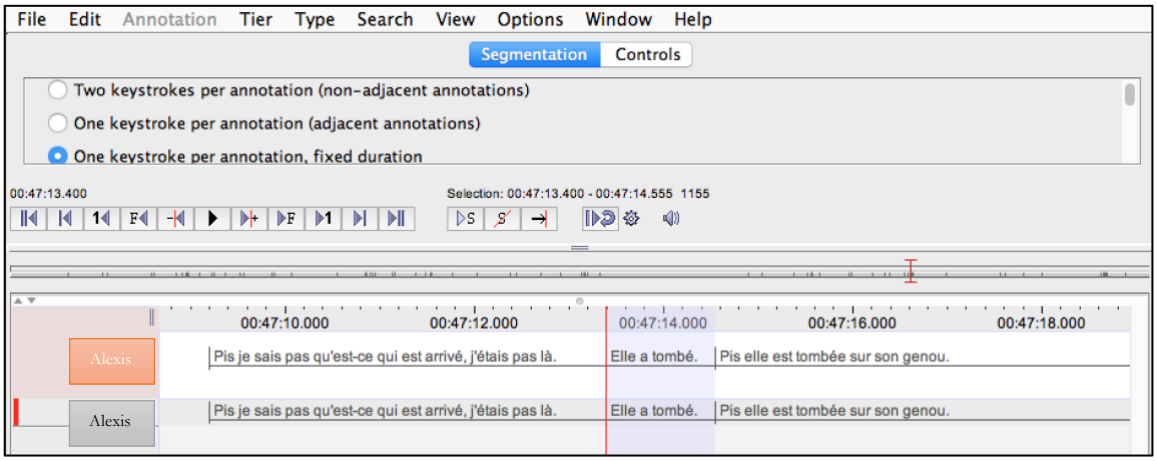

Figure 3.1 Screen capture of the segmentation mode in *ELAN*

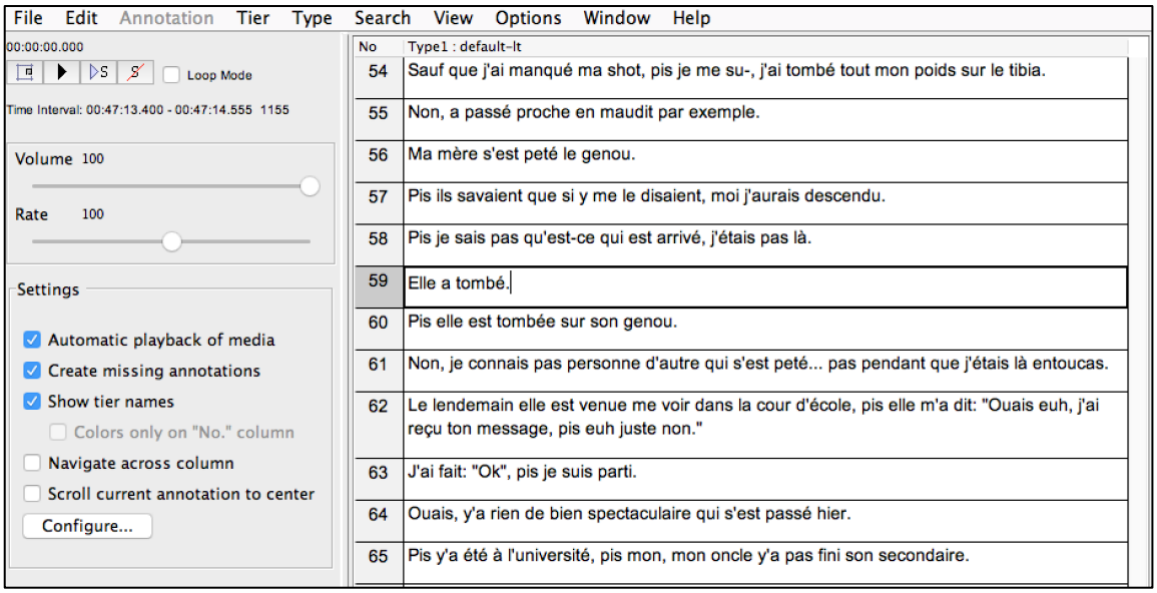

Figure 3.2 Screen capture of the transcription mode in *ELAN*

The *ELAN* software then allows the user to export the transcriptions directly into Microsoft Excel, which was then used to create a coding CSV file uploaded in *Shiny Rbrul* (Johnson 2017) for the multivariate statistical analysis.

Because of time constraints, it was thought best to segment and transcribe only the contexts where the two variants (*avoir* and *être*) appeared as auxiliaries, namely all the periphrastic tenses of the intransitive verbs under study and of the random pronominal periphrastic forms elicited during the interviews. The full utterance unit where the variant appears was first segmented and then transcribed. In the cases where a sentence consists entirely of the variant (such as *Oui, je suis resté.* 'Yes, I (have) stayed.'), then the previous sentence was transcribed as well, to provide more contextual information for the analysis.

3.4. Exclusions from the corpus

The variable data collected include most but not all the verbs studied by Sankoff & Thibault (1977): *changer*, *demeurer,* and *entrer* were excluded from my analysis. *Changer* was excluded because it was treated in Sankoff & Thibault (*ibid*: 100-101) as belonging to a separate semantic category, being the only change-of-state verb on the list, and because it is not traditionally considered an *Ê*-verb. *Changer* surfaced with *avoir* with a 0,98 probability in 1971 (*ibid*: 96). Interestingly, *changer* was not included in Willis' analysis of auxiliary alternation in Ottawa-Hull French, potentially indicating that there were no tokens of *changer* with *être* in her corpus, even though the verbs *apparaître* and *commencer* were included (Willis 2000: 52).[144]

[144] It is not currently possible to verify whether there were any tokens of *changer, apparaître,* and *commencer* in my corpus surfacing with *avoir* since tokens with these verbs have not been transcribed for this project.

Entrer and *demeurer* were excluded from my analysis because of the very small number of recorded tokens. Only two tokens of *entrer* were collected, shown in (56), both conjugated with *être*, though (56a) had a false start with *avoir* and a self-correction to *être*:

(56a) *Mon vélo* ***a, est*** *vraiment* ***entré*** *dans, dans ces poteaux-là.* (Denis)
'My bicycle really hit these poles'

(56b) *Quand on* ***est entrés*** *à l'Île-des-Soeurs, y'a un rond-point en entrant à l'Île-des-Soeurs, ...* (Sabrina)
'When we entered l'Île-des-Soeurs, there is a roundabout when you enter l'Île-des-Soeurs…'

Sankoff & Thibault (1977: 99) also had too few tokens of *entrer* for detailed discussion (17 tokens in total, with 1 token of *avoir*) and those were almost exclusively produced by highly educated speakers (Sankoff 2019: 209).[145] This is because Montréalers systematically use *rentrer* (109 tokens in total, with 68 tokens of *avoir,* in 1971).[146] In order to express iterative aspect with *(r)entrer,* they tend to use the form *re-rentrer*, as shown in (57).

(57) *T'****es****-tu* ***re-rentré*** *dans le magasin ?* (Jean-François)
'Did you go back in the store?'

Sankoff & Thibault (1977: 103) had recorded a very high rate of *avoir* selection with *demeurer* (97%, 64 tokens out of 66) in 1971, but in this study only two tokens were collected, both conjugated with *avoir*, as shown in (58). Example 58(a) is also a self-correction from *être*:

[145] In contrast, tokens in (56) were produced by one speaker who has completed university education (Sabrina) and by another who has only obtained a technical CÉGEP diploma (Denis).
[146] The two *entrer* tokens have not been merged with those of *rentrer* because since Sabrina did not produce any tokens of *rentrer,* it is impossible to verify whether she uses *rentrer* as an iterative form of *entrer*.

(58a) *On* ***est****, on* ***a demeuré*** *là trois, quatre ans il me semble.* (Annouck)
'We lived there three, four years, I think'

(58b) *On* ***a demeuré*** *là quatre ans, pis ensuite on a construit la maison pis on a emménagé.* (Denise)
'We lived there four years, and then we built the house and moved in'

Devenir, a canonical *Ê*-verb, had to be excluded from the present analysis because it did not show any alternation: all 42 tokens collected, including one of *redevenir*, were conjugated with *être*. One token of *devenir* did surface in a probable false start with *avoir* followed by a self-correction to *être*, as shown in (59).

(59) *Faque ça* ***a*** *juste, c'****est*** *juste* ***devenu*** *normal.* (Marie-Lou)
'So it just, it just became normal'

Interestingly, *(re)devenir* was not included in the 1977 Montréal analysis, for the same reason (Sankoff 2019: 203), but it was included in the Ottawa-Hull study and displayed a 12% rate of *avoir* selection (2 out of 17) (Willis 2000: 55).

3.5. Analysis of the production data

3.5.1. Apparent-time and real-time analyses

In order to minimize the interpretation and methodological pitfalls of both apparent- and real-time approaches, as described in §2.5., this study makes use of both types of analysis. It is first conducted in apparent time with two adult age groups, assuming that most features of language are acquired during childhood and teenage years and remain relatively unchanged throughout an individual's lifetime, that is to say that the speech of older generations would reflect an earlier stage of language. In the specific case of auxiliary alternation, the validity of

this assumption is confirmed by Sankoff (2019), who has revealed that auxiliary selection patterns remain stable throughout the lifespan of an individual (see sections §4.2.6. and §6.2.1.). Moreover, since the sample divisions as well as the structure and content of the sociolinguistic interview for the 2016 corpus (see §3.1.2 and §3.1.3) have been modelled as faithfully as possible on those of the 1971 and 1984 Montréal French corpora (see §2.4.2.1.), the work of Sankoff & Thibault (1977) as well as Sankoff's 2019 follow-up with the combined 1971, 1984, and 1995 panel data serve as a comparison point with which to make a real-time (or longitudinal) analysis of this linguistic variation phenomenon in Montréal French. This comparison will be crucial to verify the significance of potential generational differences that could emerge in apparent time. Additionally, since Sankoff & Thibault (1977) had not tested for *avoir* use with pronominal verbs, Gillian Sankoff revisited the 1971, 1984, and 1995 Montréal corpora in 2016 on my behalf in order to verify in what proportions pronominals selected *avoir* at the time, and these newly extracted data are used for a real-time comparison with my 2016 data in §4.3.3.

3.5.2. Variable rule analysis and *Shiny Rbrul* (Johnson 2017)

In 'real language' data, it is typical to observe an uneven distribution of categories because the number of occurrences of each context depends on its relative frequency in discourse: the number of tokens per context is variable and many combinations of variables might not be recorded (Tagliamonte 2006: 137). This is why standard statistical procedures such as ANOVA[147] are not suitable for analysis of language in use since they assume even distribution of data.[148]

[147] The analysis of variance is an estimation procedure used to analyse the differences among means in a sample, such as the 'variation' among and between groups.
[148] As Tagliamonte (2006: 137) explains, real language data "contrasts with that found in psycholinguistic studies, which are usually based on experimental data, rather than on corpus work. In this tradition the same number of examples is collected for all contexts, ensuring balanced cells and permitting analysis by standard statistical procedures."

The statistical analyses for this study were carried out with the variable rule software *Rbrul* (Johnson 2009), designed to provide a multivariate quantitative model of a situation where speakers alternate between different forms that have the same meaning in such a way that the probability of choice of either form is linguistically and/or socially conditioned. The development of the interface *Rbrul* (Johnson 2009), within the open-source statistical software *R* (*R* Core Team 2004-2019), has improved the accuracy of statistical significance. For more than 30 years, variationist studies (including every previous study on auxiliary alternation) made use of the statistical software *VARBRUL* and its successor *GoldVarb* (Cedergren & Sankoff 1974; Sankoff 1975; Rousseau & Sankoff 1978; Sankoff *et al.* 2005). However, via simulations as well as real data, Johnson (2009: 363, 376-377) shows that *GoldVarb* tends to overestimate "potentially drastically […] the statistical significance of external factors such as age and gender." This is because *GoldVarb* does not take speaker grouping into account and as a result, "GoldVarb necessarily ignores the grouping and treats each token as if it were an independent observation." (Johnson 2009: 363). In order to fix this sort of Type I error, misidentifying a chance effect as a real effect, *Rbrul* (Johnson 2009) can treat individual speakers as random effects which "takes into account that some individuals might favor a linguistic outcome while others might disfavor it, over and above (or 'under and below') what their gender, age, social class, etc. would predict" (Johnson 2009: 365).[149] *Rbrul* (Johnson 2009) can also support continuous independent variables like the age of the speakers for example (see Johnson 2009, 2014 for more details on the benefits of using *Rbrul* over *GoldVarb*).

A multivariate mixed-effects analysis of the data was performed using *Shiny Rbrul* (Johnson 2017), a version of *Rbrul* (Johnson 2009) that works through a user-friendly browser interface, illustrated in Figure 3.3 (p. 147). After uploading an Excel CSV data file on the new browser-based platform and selecting the 'response' variable (the dependent variable, i.e. *avoir*

[149] However, Johnson (2009: 365) notes that "this more conservative behavior has a trade-off: in some situations, Rbrul is more likely than GoldVarb to make a Type II error by failing to identify an effect that really does exist."

or *être*) in the left column of the top centre grey box, the user then drags each independent variable that requires testing from the white space below the box to the right column ('potential') of the box. Factor groups that have several factor levels have the option to be tested as random or fixed effects. The fixed-effects factor groups (the independent variables) that are tested by the software are each represented by differently coloured squares (categorical factor groups) and circles (continuous factor groups). Random effects are represented by empty circles. The software then automatically lists all the factor groups with a significant effect, called "potential predictors" in *Shiny Rbrul* (Johnson 2017), in decreasing order of likelihood of their having a statistically significant effect on the variation (using *p* values) in the bottom box to the right of the screen. The left of the screen lists all the factor groups that are included in the dataset, along with their individual factor levels and how many tokens they contain.

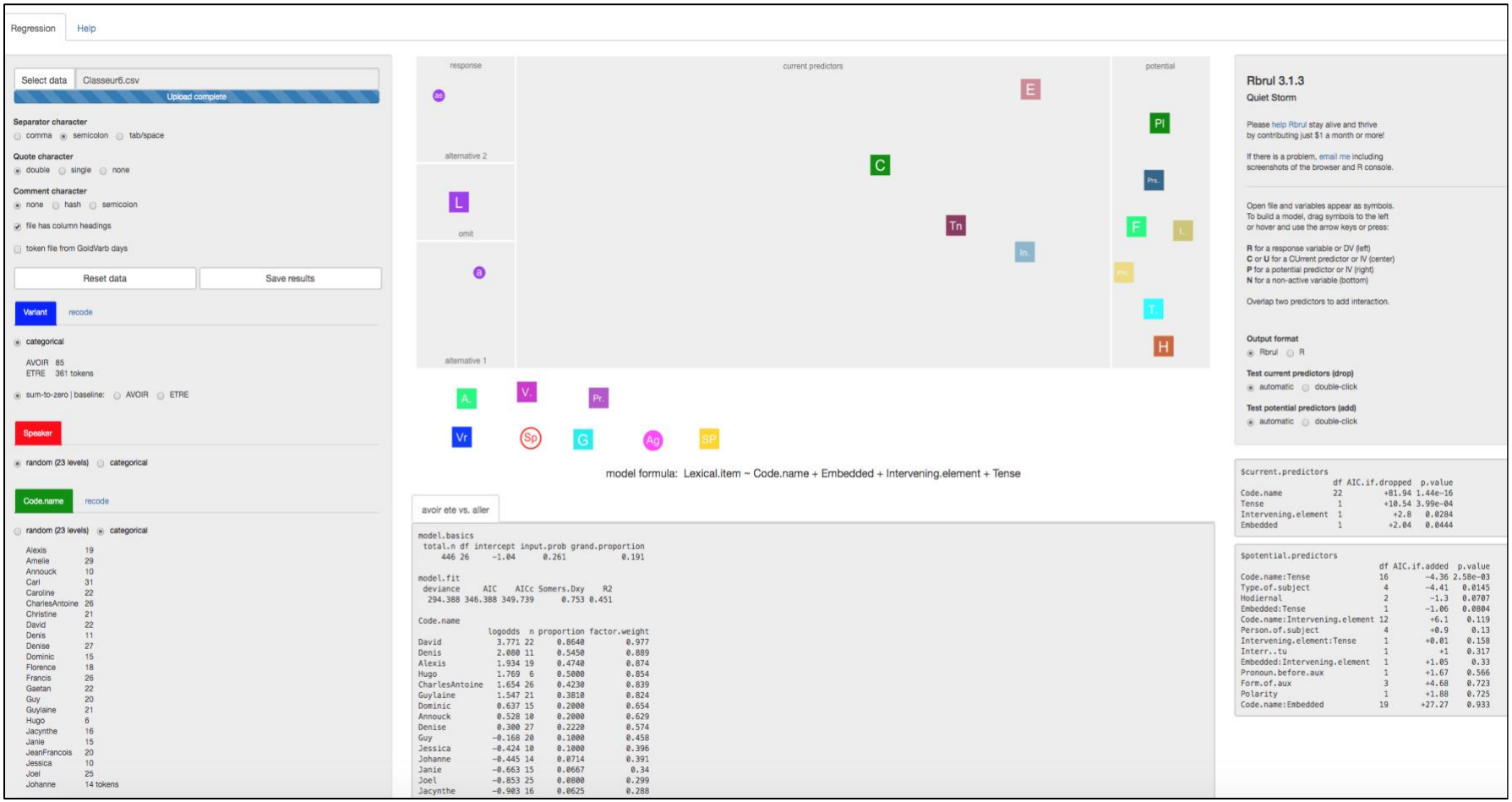

Figure 3.3 Screen capture of the *Shiny Rbrul* (Johnson 2017) browser interface.

The threshold of statistical significance is set by *Shiny Rbrul* (Johnson 2017) at $p < 0.05$, as is customary in the field.[150] The user then adds the factor groups that are predicted to have a significant effect on the response variable one by one to the model, starting from the one with the most significant effect, by dragging them from the "potential (predictors)" column (on the right of the box) to the "current predictors" one (in the middle). The software recalculates automatically the statistical impact of the remaining potentially significant factor groups as well as the new significant *p* values of the growing set of "current predictors" every time a new significant factor group is detected (in the middle box to the right of the screen). Johnson (2010, not paginated) points out that this new "interface discourages the habit of stepwise regression, a procedure that is statistically flawed and nearly taboo outside our field".

Once all factor groups with significant effects on the response variable have been added to the model, the software then lists all potential interactions between them, in decreasing order of statistical likelihood (*p* values). The best-fitting model can be identified by the user if the model stops being able to 'converge'[151] when new "predictors" are added and/or when they are no significant factor groups left to add to the model. If a model cannot converge, a message appears below the "potential predictors" box to the right of the screen. *Shiny Rbrul* (Johnson 2017) lists all total N, degrees of freedom, intercept, input probability,

[150] The levels of statistical significance that will be used in this study are as follows:

***P* value**	**Level of significance**	**Summary**
< 0.001	Very highly significant	***
0.001 to 0.01	Highly significant	**
0.01 to 0.05	Significant	*
≥ 0.05	Not significant	ns

[151] When a certain degree of accuracy is reached in one of the iterations of a variable rule analysis.

grand proportion, and indicators of model fit (deviance,[152] R-squared,[153] and AIC[154] scores, among others), as well as the individual factors of all factor groups with significant effects, in the grey rectangle at the bottom centre. The individual factors, listed per factor group, are ranked in decreasing order of factor weight[155] (or log-odds[156] for continuous factor groups, such as 'Age of the speaker'), and *Shiny Rbrul* (Johnson 2017) includes the total number of tokens per factor as well as their rate of selecting the application value, i.e. the variant defined as the outcome of the variable rule, in this case the auxiliary verb *avoir*.

3.6. Triangulation methods

Two different triangulation methods were employed in order to establish if the trend shown by the production data could actually be verified. The first triangulation method is the analysis of grammaticality judgements collected after the sociolinguistic interview and the second one is the analysis of the self-reporting judgements collected by the crowdsourcing online platform *Français de nos régions* (Avanzi *et al.* 2016).

[152] The deviance is a measure of how well the model fits the data, or how much the actual data deviate from the predictions of the model. The larger the deviance, the worse the fit. As we add predictors to the model, we will see this number decrease.

[153] R^2 is a statistical measure of how close the data are to a fitted regression line: the linear regression calculates an equation that minimizes the distance between the fitted line and all of the data points. It is the percentage of the response variable variation that is explained by a linear model: an R^2 of 1 (or 100%) indicates a perfect correlation between the predicted model and the data, i.e. that the model explains all the variability of the response data around its mean.

[154] The Akaike Information Criterion (AIC) provides a method for assessing the quality of the model through comparison of related models. It is based on the deviance, but "penalizes you for making the model more complicated. Much like adjusted R-squared, its intent is to prevent you from including irrelevant predictors. However, unlike adjusted R-squared, the number itself is not meaningful. If you have more than one similar candidate models (where all of the variables of the simpler model occur in the more complex models), then you should select the model that has the smallest AIC. It is useful for comparing models, but is not interpretable on its own." https://www.theanalysisfactor.com/r-glm-model-fit/ (webpage visited on February 20, 2020)

[155] Factor weights indicate whether a given factor favours a specific variant (> 0.5), "is likely to occur" or disfavours it (< 0.5), "is not likely to occur".

[156] The log-odds (or logit) is the logarithm of the odds $\frac{p}{1-p}$ where p is probability. In the case of age, for example, it could be the case that with each additional year the probability of a specific outcome increases.

3.6.1. Grammaticality judgements data

After the sociolinguistic interview, the 48 speakers performed a grammaticality judgement task on the acceptability of auxiliary *avoir* in sentences with the 16 verbs studied, as well as with various pronominal verbs. This study aims to discover whether people who do not really use the *avoir* variant in their own speech still recognize that it can be found in the local variety. I also seek to establish a hierarchy of which verbs people think are more or less acceptable with *avoir* than others and correlate these results with the interview data.

The participants were asked to listen to 60 short sentences[157] being read aloud to them, and to say whether they could have potentially heard these sentences, some of them containing the variable under investigation, being uttered 'naturally' by native speakers (see Appendix 3C for the entire list of prompts with their English translation). The speakers were not told prior to the exercise that their grammaticality judgements were collected to study a specific linguistic variable, i.e. auxiliary alternation. Since this morphosyntactic variable is strongly correlated with prescriptive grammar, as opposed to a phonological variant, asking the subjects to say whether they personally would say sentences with *avoir* or to perform an elicitation task would have been problematic because it is likely that most of them would deny using the *avoir* variant or would produce a vast majority of *être* auxiliary tokens, even if they had been recorded producing tokens with *avoir* during the sociolinguistic interview.

All prompts were read by me[158] rather than directly by the participants in order to neutralize the influence of the written language on their answer, assuming that reading a non-standard form might prompt the speakers to realize that it is a grammar "mistake" and that it is "ungrammatical". However, if they could only hear the token, and additionally in some cases focused their answer on the content of the sentence rather than on its structure, it would

[157] The order in which the sentences were read was the same for all the participants. However, the initial order of the sentences was established randomly. There is therefore no way of knowing whether the sentence order has conditioned the answers of the participants.
[158] Future studies including a grammaticality judgement task could make use of recorded prompts played to the participants in order to ensure consistency across sessions.

mean that they had accepted the variant. Their grammaticality judgements therefore concerned whatever they thought was relevant. There are consequently multiple possible variables that I have not been able to control for in the answers, such as the influence of the semantic content, the appropriateness of the register, the grammaticality of another part of the sentence, and the co-occurrence of other non-standard features.

In the grammaticality judgement test, each *É*-verb was tested with two prompts: one containing the *avoir* variant and one 'control' sentence containing auxiliary *être*, but it should be noted that the two sentences with the same verb were different in their content as well, so as not to alert the subject to the variable under study. Most of the *avoir* examples were taken from the pilot study interview recordings (Rea 2014: 110-115), except sentences with *aller, naître, parvenir, devenir, mourir, venir, survenir,* and *intervenir,* so that the 48 subjects would be judging real-language-in-use data. A few *avoir* examples are presented in (60).

(60a) *J'**ai sorti** avec elle pendant deux ans à peu près.*
'I dated her for about two years.'

(60b) *On **a resté** neuf jours à Paris, puis après on a pris le train pour Londres.*
'We stayed nine days in Paris, then after that we took the train to London.'

(60c) *On **a retourné** en arrière parce qu'on avait oublié le chien à la maison.*
'We turned back because we had forgotten the dog at home.'

(60d) *Je **m'avais déguisé** en Père Noël pour les enfants.*
'I had dressed up as Santa Claus for the children.'

However, because of this methodology, it was not possible to control for the tense or the person/number of the subject of the sentences, nor for their content.

In §5.1.1., I seek to determine whether the ranking of the various *É*-verbs in terms of *avoir* acceptability rate is the same as the *avoir* selection with the production data, and whether it is possible to correlate the two rates with a scatterplot fitted with a regression line (giving

an R^2 score[159]). I also verify whether there is a discrepancy between the 48 individual speakers' willingness to accept intransitive and pronominal verbs that surface with auxiliary *avoir* and the individual speakers' actual use of *avoir* in the interviews, and whether it is again possible to correlate the two rates with a scatterplot fitted with a regression line (§5.1.2.). Lastly, I examine whether social variables can be held responsible for the disparity among the 48 individual speakers in *avoir* acceptability rates recorded in intransitive and pronominal sentences during the grammaticality judgement task (§5.1.3.). In order to test for this correlation, I sort the speakers into their respective social categories (SPS levels, age, etc.) and average their acceptability rates by sub-category (LOW SPS, MID SPS, HIGH SPS, young, old, etc.). I then compare whether the ranking of constraints within each social variable is the same for the judgement data and the production data collected during the interviews.

3.6.2. The crowdsourcing platform *Français de nos régions* (Avanzi *et al.* 2016)

The blog *Français de nos régions*[160] (Avanzi *et al.* 2016; for the epistemological aspects of the project, see Avanzi & Thibault 2018) launched by Mathieu Avanzi and André Thibault in 2015, aims to map linguistic variation through postal codes, by using crowdsourcing surveys to track various lexical, syntactic and phonological usages across the French-speaking world. The survey also collects the metalinguistic information of participants, such as the gender, age, last level of education completed, etc. In 2017, Avanzi & Thibault collected 3748 native grammaticality judgements in North America and Europe on the two pairs of written sentences found in (61a) and (61b), respectively an intransitive one and a pronominal one. The survey asked whether the participants would say each sentence with *être,* with *avoir* or with both indiscriminately.

[159] See footnote 153 for a definition of the R^2 score.
[160] See https://francaisdenosregions.com/

(61a) *Je suis/J'ai monté sur le toit de la maison*
'I have climbed on the roof of the house'

(61b) *Je me suis/J'm'ai lavé les mains*[161]
'I have washed my hands'.

Out of the 3748 answers collected, 821 participants lived in the Greater Montréal[162] area. The Greater Montréal sample comprises 590 women and 219 men (12 preferred not to reveal their gender). The corpus is also unbalanced in terms of education levels: 1 participant had only completed elementary school, 33 secondary or professional school, 167 had attended CÉGEP or an equivalent, and 620 had had university-level education. The age of the participants ranged from 16 to 79 years old.

In §5.2., I run two multivariate statistical analyses with *Shiny Rbrul* (Johnson 2017) in order to test whether the social conditioning for *avoir* use with intransitive and pronominal verbs observed in my production data is also at play with regard to the *avoir* acceptability rate with these two pairs of sentences. I followed the same regression steps as those outlined in §3.5.2. but with only four independent social variables or "potential predictors" (gender, age, level of education, and municipality of the participant). As previously stated, the participants had the option to indicate whether they would say the two pairs of sentences either with *être,* with *avoir* or with both auxiliary verbs indiscriminately, but while the software *Shiny Rbrul* (Johnson 2017) can now analyse tripartite variables, it was thought best for comparison purposes and since the data deals with self-reports, to collapse the tokens of the participants

[161] It is not entirely clear to me why Avanzi & Thibault decided to use a more 'oral' spelling, i.e. the contraction of the subject pronoun with the object pronoun, in the *avoir* variant than in the *être* variant of the pronominal sentence. It is likely that this difference in spelling has impacted the variant choice of the participants of the survey.

[162] Including 450 from the Island of Montréal, and 68 from Laval, 63 from La Vallée du Richelieu, 40 from Roussillon, 34 from L'Assomption, 33 from Marguerite D'Youville-Lajemmerais, 33 from Sainte-Thérèse-De Blainville, 31 from Vaudreuil-Soulanges, 25 from Deux-Montagnes, 24 from Les Moulins, 11 from Beauharnois-Salaberry, as well as 9 from Mirabel.

who chose the *avoir* variant with those of the participants who indicated that the sentence was felicitous regardless of the auxiliary verb.

In order to perform a multivariate analysis on the corpus to determine which social factors might have influenced the distribution of the data, the only individual with no more than elementary education and the 12 tokens of the individuals without an assigned gender were excluded from the analysis. After removing these individuals from the sample, I obtained 808 judgements for the intransitive sentence and 805 for the pronominal one.[163] All of the judgements of the participants who were not from the Island of Montréal were collapsed into one factor group, to investigate whether there would be geographical differences between the centre and the periphery of the Greater Montréal area. There was insufficient data from each municipality for a more fine-grained geographical analysis.

[163] In the pronominal verb section of the survey, three participants did not provide an answer.

Chapter 4. Analysis of the production data

4.1. Introduction

The research questions (see §2.6. above) examined here are 1) whether there have been changes in auxiliary selection rates in Montréal French since 1971 (as well as since 1984 and 1995), and, if so, which direction the observed changes have taken; 2) whether there have been changes in the social and linguistic conditioning of the variable, and if so, which changes have taken place; 3) whether these intransitive and pronominal data turn out to be distributed in similar social and linguistic patterns to those that have already been recorded in other French and Romance varieties (see Chapter 2); and finally 4) whether it is possible to detect auxiliary alternation in pronominal verbs.

This chapter will first lay out the overall distribution of auxiliaries in intransitive verbs (§4.2.1.). Second, it will present a series of multivariate analyses in order to determine which predictors can most accurately explain the auxiliary alternation recorded in the data (§4.2.2.). Third, I perform distributional analyses with descriptive statistics of social and linguistic variables that were tested as random (§4.2.3.) and fixed effects (§4.2.4.) in the mixed-effects models, and that could not be included in the multivariate analyses (§4.2.5.). These results will then be compared and contrasted with Sankoff & Thibault's 1977 study and Sankoff's 2019 update (§4.2.6.). The chapter will then analyse auxiliary alternation that surfaced with pronominal verbs in the 2016 data (§4.3.1. for overall rates and §4.3.2. for the detailed distribution) and will subsequently compare such results with a preliminary analysis of auxiliary alternation in pronominal verbs in the 1971, the 1984, and the 1995 corpora (§4.3.3.).

In §4.4., I will analyse two types of possible 'avoidance mechanisms', i.e. instances where speakers use various linguistic strategies equivalent in meaning but different in form from periphrastic tenses with auxiliary *être*. The first one only concerns the verb *aller*: using

avoir été in place of *être allé* (§4.4.1.) and the second one only concerns the *passé composé* tense: the use of historic present (§4.4.2.).

4.2. Intransitive verbs

4.2.1. Overall variant distribution

The proportion of non-standard auxiliary selection with intransitive verbs has fallen by two thirds since 1971, but a substantial number remain in the 2016 data, as shown in Table 4.1 and illustrated in examples 62-64 (all extracted from the 2016 corpus), each showing the same verb being used with both *avoir* and *être*.

	Corpus 1971		**Corpus 2016**	
Variant	**%**	**Total N**	**%**	**Total N**
Avoir	32,8	719	10,7	253
Être	67,2	1474	89,3	2121
Total		2193		2374

Table 4.1 Distribution of auxiliaries in periphrastic tenses of intransitive *Ê*-verbs in Montréal French in 1971 and in 2016

(62a) *Mon frère, il descendait une côte pis sa roue d'en avant* ***a parti.*** (Rachel)
'My brother, he was going down a hill and his front wheel came off'

(62b) *Moi, je* ***suis partie*** *de chez nous à 17 ans là.* (Sara)
'I left home at 17 years old [discourse particle]'

(63a) *On* ***a*** *juste* ***resté*** *là, pis on a bu, pis on se promenait pis c'était le fun.* (Charles-Antoine)
'We just stayed there, and we drank, and we were walking around and it was fun'

(63b) *Je sais pas comment ça se fait, l'eau* ***était restée*** *chaude.* (Annouck)
'I don't know how that happened, the water had stayed warm'

(64a) *Mais après une semaine, j'ai, j'****ai retourné*** *travailler.* (Gaëtan)
'But after a week, I, I went back to work'

(64b) *Après ça on* ***est retournés*** *à l'appartement, j'avais mal aux pieds.* (Sabrina)
'After that we came back to the apartment, my feet were hurting'

4.2.2. Multivariate analyses with *Shiny Rbrul* (Johnson 2017)

4.2.2.1. Methodology

As outlined in Chapter 3, the factor groups tested by the variable rule analysis are both linguistic ((in)animacy of the subject, hodiernal action, presence of a clitic pronoun before the auxiliary, intervening material between the auxiliary and the past participle, possibility of transitive use, possibility of pronominal use, tense of the verb, polarity of the sentence, type of subject, clause structure, form of auxiliary, person of subject, morphologically-derived verb) as well as social (age, gender, SPS, contact with English, place of childhood).

In order for the statistical model to converge, when the tokens of factor levels within factor groups were too few, they were removed from the analysis so that the data was as evenly distributed as possible. This is notably the case with 2PL subjects (7 tokens), some types of subjects (5 tokens of *qui est-ce qui* 'who', 3 of *quelqu'un* 'someone', and 1 of *tout* 'everything'), past subjunctives (12 tokens), and past infinitives (17 tokens). From the remaining tokens, all false starts with *avoir* followed by a self-correction to *être* (33 tokens) were also excluded from the main analyses because of how much they can vary in terms of material surfacing in the false start and/or in the self-correction. The total number of tokens analysed with *Shiny Rbrul* (Johnson 2017) therefore initially went down from 2374 to 2309 tokens for the model for comparison with the earlier studies (model presented in §4.2.2.2.). However, some common verbs, such as *aller, arriver, (re)partir,* and *(re)venir*, select *avoir* very infrequently, and while mixed-effects modelling should be able to cope with such unbalanced data, these few verbs make up practically two-thirds of my data. I therefore performed a second analysis without these frequent verbs in order to obtain a better-fitting model (presented in §4.2.2.3.). After removing such near-categorical items, the total number of tokens went down from 2309 to 683 tokens.

Moreover, other factors were collapsed in order for the subsequent models to converge: lexical verbs and their iterative forms (from 27 to 13 verbs), as well as hodiernal tokens and the 14 tokens in which the action described was thought by the participants to be hodiernal but had actually taken place within 48 hours (see §3.2.3.15.).

In order to account for the fact that speakers use varying lexical items (and do so in different proportions), and that these various verbs select different auxiliaries according to a number of factors, the effects of both the speaker and the lexical item were set as random in *Shiny Rbrul* (Johnson 2017). The first full comparative model, Model 1 (which is given in Appendix 4A), was produced with three separate socioprofessional statuses, but would not converge properly. It transpired that the factor weights for the MID and HIGH socioprofessional statuses were almost identical (0.402 and 0.441 respectively), so they were collapsed into one group opposed to LOW SPS (which yielded a FW of 0.653). It might be the case that the similarity between the MID and HIGH socioprofessional statuses, in terms of FWs, is the result of a methodological issue (for example how the speakers were chosen and then assigned to a category) or is indeed the result of an actual quasi merger of both classes in terms of how they impact this variable.

4.2.2.2. Comparative model

The mixed-effects model presented in Table 4.2, Model 2, includes all the verbs studied by Sankoff & Thibault (1977) that also showed alternation in my data, even at very small rates, in order to provide a direct comparison with the data collected in 1971. The table ranks significant factor group effects in decreasing order of statistical significance (by p value) where the threshold of statistical significance has been set by *Shiny Rbrul* (Johnson 2017) at $p < 0.05$, as is customary in the field. As stated in §3.5.2., statistical significance is reached when the variable rule program assesses that the results were not produced by chance. The table also

presents in decreasing order the factor weights of the individual factor levels (within each significant factor group). The factor weights calculate the impact that an independent variable has on a dependent one and indicate whether a given factor favours a specific outcome (> 0.5), i.e. "*avoir* is likely to occur", has no effect on the outcome (= 0.5), or disfavours it (< 0.5), i.e. "*avoir* is not likely to occur". The total number of tokens per factor level is also included in the table as well as their rate (in %) of selecting the application value, i.e. the variant defined as the outcome of the variable rule – in this case the auxiliary verb *avoir*. No significant interactions were identified.

MODEL 2	**Multivariate analysis of the likelihood of *avoir* selection (vs *être*) with two SPS levels**		
Model formula: Variant ~ (In)animacy of the subject + Hodiernal.recoded + Intervening.element +Pronoun.before.aux + SPS.recoded + Tense + Transitive.use + (1 | Code.name) + (1 |Verb.collapsed)			
Input probability	0.306		
Total rate	10,8%		
Total N	2309		
AIC	895.333		
R^2	0.682		
Deviance	873.333		
Significant factor groups[164]	**Factor weights**	**% *avoir***	**Total N**
(In)animacy of subject* **	p = 4.97e-07		
Inanimate	0.67	23,4	214
Animate	0.33	9,5	2095
Pronoun before auxiliary **	p = 0.00166		
Pronoun before aux	0.718	20	55
No pronoun before aux	0.282	10,6	2254
Socioprofessional status **	p = 0.00352		
Low	0.617	14,9	698
Mid-High	0.383	9	1611
Hodiernal action **	p = 0.00692		
Pre-hodiernal	0.614	12,3	1909
Hodiernal	0.386	3,5	400
**Intervening element ** **	p = 0.00859		
Intervening element	0.597	18,1	182
No intervening element	0.403	10,2	2127
Tense*	p = 0.0115		
Conditionnel passé	0.749	28	25
Passé composé	0.477	10,9	2109
Plus-que-parfait	0.268	7,4	175
Transitive use*	p = 0.0118		
Transitive use allowed	0.826	24,8	734
Transitive use not allowed	0.174	4,3	1575
Speaker (48)	Random		
Lexical item (13)	Random		

Table 4.2 *Shiny Rbrul* (Johnson 2017) results for Model 2: statistical effects of the significant factor groups on auxiliary alternation in the 2016 Montréal data, by p values and factor weights

[164] The factor groups which did not prove to have a significant effect are 'Gender of the speaker', 'Bilingualism', 'Type of subject', 'Clause structure', 'Form of auxiliary', 'Person of subject', 'Morphologically-derived verb', 'Age of the speaker (continuous)', 'Age of the speaker (categorical = young vs old)', 'Possibility of pronominal use'.

Table 4.2 shows an input probability[165] (or corrected mean) of 0.306, indicating that *avoir* has a 30,6% chance of surfacing, even though *avoir* surfaced in my corpus only in 10,8% of cases. The predictors of *avoir* selection are mostly linguistic apart from the socioprofessional status of the speaker. The factor group that was identified by *Shiny Rbrul* (Johnson 2017) as the one with the strongest effect, a very highly significant one (p = 4.97e-07), is the (in)animacy of the subject: inanimate subjects favour *avoir* with a FW of 0.67 (the animate subjects disfavour *avoir*, with a FW of 0.33). Inanimate subjects were ten times less frequent than animate ones but surfaced with *avoir* three times as frequently. Example tokens can be found in §3.2.3.11. As mentioned in §2.4.2.3.11., this factor group was tested by Sankoff (2019: 223-224) on the combined 1971, 1984, and 1995 Montréal data, and my results are consistent with hers, i.e. inanimate subjects favoured *avoir* selection in her data as well (p < 0.01). Willis (2000: 66) also tested for this factor in her Ottawa-Hull study but did not find it to be significant.

The factor group with the second strongest effect in my data is whether a pronominal clitic precedes the auxiliary verb. This effect is highly significant (with a p = 0.00166). The presence of a pronominal clitic preceding the auxiliary verb favours *avoir*, with a FW of 0.718, and the absence of a pronominal clitic disfavours it, with a FW of 0.282. Example tokens can be found in §3.2.3.17. This is the first time that such factor group is tested on auxiliary alternation data and it was included after careful examination of all *avoir* tokens collected. Although this factor group has not been tested before, the fact that it has the second strongest p suggests that it should have been.

The factor group with the third strongest effect is the SPS, with a highly significant p = 0.00352. The membership of the lowest socioprofessional status yielded a FW of 0.617, favouring *avoir* (and the tokens which were uttered by the MID-HIGH speakers had a FW of 0.383, disfavouring *avoir*). This result does not come as a surprise, since the vast

[165] The likelihood that a speaker will choose the application value, here auxiliary *avoir*, regardless of the presence or absence of any other factor in the environment (Bayley *et al.* 2013: 18).

majority of variationist studies mentioned in §2.4.2. have correlated lower socioprofessional status with higher *avoir* selection than the other socioprofessional categories. As mentioned in §2.4.2.2.3., these factor groups were re-tested by Sankoff (2019: 223-224) on the combined 1971, 1984, and 1995 Montréal data, and my results again are consistent with hers, i.e. a low SPS (as well as the least educated speakers[166]) favoured *avoir* selection in her data as well ($p < 0.001$).

The factor group with the fourth strongest effect, whether the action reported by the speakers was hodiernal, i.e. had taken place within 24 hours of (or the day prior to) the moment of speech, has a highly significant $p = 0.00692$. When the action reported by the speakers was pre-hodiernal it yielded a factor weigh of 0.614, favouring *avoir* (and the actions which were hodiernal had a FW of 0.386, disfavouring *avoir*). Example tokens of hodiernal and pre-hodiernal tokens were provided in §3.2.3.15. This factor might be connected to the attention that the speakers paid to their speech: every interview included a question asking the participants to recount in detail what they had done the previous day.[167] As stated in §3.1.3., the goal of adding such a question was to collect a maximum number of tokens with pronominal verbs, since many are related to daily routines and personal grooming, as well as to test the hodiernal hypothesis. However, since such a question required the participants to make an effort in remembering specific events in a specific order, as opposed to talking freely, I noticed that their speech was consequently slower. While this fact does not necessarily entail that the participants were paying more attention to their speech, psycholinguistic and/or prosodic analyses might be more appropriate tools to reveal the complex mechanisms behind this predictor and how they might have impacted auxiliary choice.

[166] As mentioned in §3.2.2.4., my socioprofessional index included in its score calculation the level of education as well as the type of occupation of the speakers.
[167] Five speakers ended up describing how they had spent the day preceding the 'previous day' or the day of the interview.

The factor group with the fifth strongest effect is whether there is intervening material between the auxiliary and the past participle. This effect is highly significant, with a $p = 0.00859$. The presence of an intervening element between the auxiliary and the past participle yielded a FW of 0.597, favouring *avoir* (and the tokens which did not have intervening material had a FW of 0.403, disfavouring *avoir*). Example tokens have been provided in §3.2.3.13. Negation adverbs were also included in this factor group, but polarity of the sentence was not found to be a significant factor of influence on the data. As mentioned in §2.4.2.3.12., this factor group was tested by Sankoff (2019: 223-224) on the combined 1971, 1984, and 1995 Montréal data, and my results are consistent with hers, i.e. intervening material (usually adverbial) favoured *avoir* selection in her data as well ($p < 0.01$).

The tense of the verb is the factor group with the sixth strongest effect identified by *Shiny Rbrul* (Johnson 2017) with a significant $p = 0.0115$. After the past subjunctive and the past infinitive tokens were removed from the analysis, the mixed-effects model shows that tense also has a small but significant effect on the alternation, with the past conditional favouring *avoir* (FW of 0.749), the *passé composé* producing almost no effect on the alternation (FW of 0.477), and the *plus-que-parfait* disfavouring *avoir* selection (FW of 0.268). It might be interesting to note that, although they have been removed from the multivariate analysis because they were too few of them, past subjunctive tokens selected *avoir* with a rate of 10% (1/10) and past infinitive ones did so with a very high rate of *avoir* selection (5/11, including 2 tokens of *avoir été* replacing *être allé*[168]), as illustrated in examples 65(a-c):

[168] See §4.4.1. for a detailed discussion on the rates of replacement of the periphrastic past tenses of *aller* with *être*.

(65a) *Est-ce que cette personne-là doit* ***avoir resté*** *à Montréal toujours toujours toujours ?* (Sylvain)
'Must this person have lived in Montréal always always always?'

(65b) *À 35 ans là je venais d'avoir N. là... je suis ben contente d'****avoir passé*** *par là.* (Annouck)
'At 35 years old I had just had N.… I'm very happy to have been through that'

(65c) *Pis* ***avoir été*** *souper chez le, son, ton beau-père, pis nous on a, c'est des, nos propres expressions, mais pis là ils nous ont fait à souper pis là j'ai dit : « Ah c'est écœurant ! »* (Jacynthe)
'And having gone for dinner at the, his, your father-in-law's, and we have, they are, our own sayings, but then they made us dinner and then I said: "Ah! It's delicious[169]/revolting!"'

The opposite effects exhibited by the *plus-que-parfait* and *conditionnel passé* are difficult to explain, but one could postulate that this difference in conditioning could be related to their different likelihoods of surfacing in different semantic or syntactic contexts, even though such an interaction was not detected by *Shiny Rbrul* (Johnson 2017). As mentioned in §2.4.2.3.9., the factor group 'Tense' was tested by Sankoff (2019: 223-224) on the combined 1971, 1984, and 1995 Montréal data, but my results are only partially consistent with hers: in her data, all tenses[170] other than the *passé composé* favoured *avoir* selection ($p < 0.05$), indicating opposite effects for the *plus-que-parfait* when compared to 2016.

The factor group with the seventh strongest effect is whether the verb can be used transitively. It has a significant $p = 0.0118$. The possibility of transitive use favours *avoir* selection in my data with a FW of 0.826, while the non-possibility of transitive use disfavours *avoir* with a FW of 0.174. Example tokens of these two factors have been provided in §3.2.3.5. The fact that the factor group 'Transitive use' is identified by *Shiny Rbrul* (Johnson 2017) as having a significant effect confirms the findings not only of Sankoff & Thibault (1977), but also of Canale, Mougeon, & Bélanger (1978), Russo & Roberts (1999), Willis (2000), Stelling

[169] In standard French, the adjective *écœurant* means 'sickening, nauseating, revolting'. However, in Québécois French, it is almost always used with a positive connotation, meaning 'delicious' or 'awesome'. Hence the confusion mentioned by speaker Jacynthe in example 65(c). However, when used as a substantive, *un écœurant* refers exclusively to a very bad person.
[170] Sankoff (2019: 207) only mentions the *futur antérieur* 'future perfect' and the *plus-que-parfait* 'past perfect'.

(2011), and Roussel (2016) (see §2.4.2.3.4.), for whom the possibility of transitive use always contributed positively to the probability of *avoir* selection.

4.2.2.2.1. Other factor groups with potentially significant effects

I will now report factor groups that are listed as 'potential predictors' in the multivariate analysis, i.e. they were identified by *Shiny Rbrul* (Johnson 2017) as factor groups with potentially statistically significant effects, but when they were separately added to the model, one by one, the model would not converge.[171] In Table 4.3, the *p* value indicated next to each factor group is the predicted *p* value of that factor group if it had been the only one added to the model.

Factor groups listed as having potentially significant effects and when individually added to Model 2 prevent it from converging			
	***p* values**		
Potentially significant factor groups	**Factor weights**	**% *avoir***	**Total N**
Polarity of the sentence	$p = 0.0327$		
Positive	0.648	10,9	2221
Negative	0.352	9,1	88
Place of childhood	$p = 0.0374$		
Outside of the Greater Montréal Area	0.586	13,1	505
In the Greater Montréal Area	0.414	10,1	1804

Table 4.3 Factor groups that are identified as potentially significant for Model 2 in *Shiny Rbrul* (Johnson 2017) but when they were individually added as current predictors the model could not converge

[171] Johnson has suggested that in these cases there would not be "a great reason to exclude [such factor group] just because the model with [such factor group] doesn't converge, but it's not ideal [...] if a model has not converged (properly), then you are on shaky ground trusting its parameters or the inferences made when comparing it to another model..." (Johnson 2019, personal communication)

The factor group 'Polarity of the sentence' has a significant $p = 0.0327$. A positive sentence favours *avoir*, with a FW of 0.648, and a negative one disfavours it, with a FW of 0.352. Example tokens of these two factors have been provided in §3.2.3.14. This finding is particularly interesting because adverbs of negation were included in the coding of the factor level 'Presence of intervening material between the auxiliary and the past participle', which favoured *avoir,* as illustrated in Table 4.2. Negative sentences, despite containing such intervening material (negative adverb *pas*), disfavour *avoir* selection.

As mentioned in §3.1.1., when trying to find native speakers of Montréal French, it was difficult to find participants who were born and raised in Montréal and were still living in the *agglomération* 'urban area'. Since speakers who have not spent their entire childhood in Montréal have been included in the corpus, it was thought best to check whether that could have an influence on auxiliary selection. The distribution of speakers was not balanced, however: there were four times more LOW SPS and MID SPS speakers who had spent their childhood in the Greater Montréal area (13 speakers each) than had been raised elsewhere in Québec (3 speakers each), as is shown in Table 4.4.

Place of childhood	LOW SPS speakers	MID SPS speakers	HIGH SPS speakers	Total speakers
Childhood in the Greater Montréal	13	13	11	37
Childhood elsewhere in Québec	3	3	5	11

Table 4.4 Distribution of speakers per SPS based on where they spent their childhood

Despite there being more than four times the number of LOW SPS speakers in the corpus that were raised within the Greater Montréal Area than those raised elsewhere in Québec, the factor group 'Place of childhood' has a potentially significant effect, with a

p = 0.0374. Speaker who were raised outside of the Greater Montréal Area tended to favour *avoir* (with a FW of 0.586) in comparison to speakers who had spent their childhood within the Greater Montréal Area (disfavouring *avoir,* with a FW of 0.414). Presuming that auxiliation patterns are stable throughout the lifespan of an individual (see §4.2.6. and §6.2.1. below), this might suggest that *avoir* use is more prominent outside the Greater Montréal Area.

4.2.2.3. Best-fitting model

I performed another iteration of Model 2, the comparative model, after removing (near-) categorical items, sometimes called 'KnockOuts' (Tagliamonte 2006: 222), i.e. lexical items or speakers which never or almost never selected *avoir*.[172] This was done in order to test whether factor groups that were selected as having significant effects in Model 2 can explain the distribution of data a lot more variable in its auxiliary selection. No other study of auxiliary alternation mentioned in §2.4.2. had removed such items from their datasets prior to statistical analysis, hence my presentation of both Model 2 and Model 3. The verbs which were excluded from Model 3 are *aller* (0,4% of *avoir* selection)*, (re)venir* (0,6%), *arriver* (0,6%), as well as *(re)partir* (2,7%). The two speakers who were also categorical users of *être*, Francis and Marie-Lou, were also excluded from the analysis. After performing these exclusions, the factor group 'Type of subject' contained three cells that were almost empty (1 token of impersonal *il*, 1 token of prodrop, and 3 tokens of proper nouns) and these were consequently also excluded from the analysis. The new dataset is therefore considerably reduced in size, with a total of 683 tokens. Model 3 is presented in Table 4.5.

[172] There were no lexical items or speakers that almost never selected *être*.

MODEL 3	**Multivariate analysis of the likelihood of *avoir* selection (vs *être*) without categorical and near-categorical *Ê*-items**		
Model formula: Variant ~ (In)animacy of the subject + Hodiernal.recoded + Intervening.element + Pronoun.before.aux + SPS.recoded + Tense + (1 | Code.name) + (1 | Verb.collapsed)			
Input probability	0.854		
Total rate	34%		
Total N	683		
AIC	709.544		
R^2	0.458		
Deviance	689.544		
Significant factor groups[173]	**Factor weights**	**% *avoir***	**Total N**
Pronoun before auxiliary* **	p = 8.31e-06		
Pronoun before aux	0.871	91,7	12
No pronoun before aux	0.129	32,9	671
(In)animacy of subject* **	p = 2.57e-05		
Inanimate	0.659	52,5	80
Animate	0.341	31,5	603
Socioprofessional status **	p = 0.00245		
Low	0.628	43,9	223
Mid-High	0.372	29,1	460
Hodiernal action **	p = 0.00464		
Pre-hodiernal	0.631	36,3	606
Hodiernal	0.369	15,6	77
Intervening element *	p = 0.0125		
Intervening element	0.6	45,5	66
No intervening element	0.4	32,7	617
Tense*	p = 0.0255		
Conditionnel passé	0.72	54,5	11
Passé composé	0.506	34,6	621
Plus-que-parfait	0.275	21,6	51
Speaker (46)	Random		
Lexical item (9)	Random		

Table 4.5 *Shiny Rbrul* (Johnson 2017) results for Model 3: statistical effects of the significant factor groups on auxiliary alternation in the 2016 Montréal data, by p values and factor weights, after the removal of categorical and near-categorical items

[173] The factor groups which did not prove to have a significant effect are 'Gender of the speaker', 'Bilingualism', 'Place of childhood', 'Clause structure', 'Form of the auxiliary', 'Pronominal use', 'Transitive use', 'Person of subject', 'Morphologically-derived verb', 'Age of the speaker (continuous)', 'Age of the speaker (categorical = young vs old)'.

Table 4.5 shows a total rate of *avoir* of 34% (compared to 10,8% in Model 2) and an input probability of 0.854, indicating that removing all (near-)categorical items almost triples the chances of *avoir* surfacing in the data. Model 3 demonstrates that practically the same social and linguistic conditioning is at play in both datasets and that the ranking of constraints is furthermore almost identical. Auxiliary *avoir* is still favoured when a clitic pronoun precedes the auxiliary verb (the factor group 'Pronoun before auxiliary' has now a very highly significant effect, with a p = 8.31e-06). It is also still favoured when the subject is inanimate (the factor group '(In)animacy of the subject' has a very highly significant effect, with a p = 2.57e-05). *Avoir* is still more likely to surface when the speaker belongs to the lowest socioprofessional status (the factor group 'SPS' has a highly significant effect, with a p = 0.00245). Moreover, *avoir* is still favoured when the action described has not taken place in the last 24 hours (the factor group 'Hodiernal' has a highly significant effect, with a p = 0.00464), when there is an intervening element between the auxiliary verb and the past participle (the factor group 'Intervening element' now has a significant effect, with a p = 0.0125, as opposed to a highly significant one in Model 2 with a p = 0.00859), and when the verb is conjugated in *conditionnel passé* (the factor group 'Tense' has a significant effect, with a p = 0.0255). The two models differ in that Model 3 fails to identify the possibility of transitive use as a factor group with a significant effect. This result was expected since among the *Ê*-verbs excluded from Model 3 (*aller*, *(re)venir, arriver*, and *(re)partir*) only *(re)partir* allows transitive use, as mentioned in §3.2.3.5.

As with Model 2, *Shiny Rbrul* (Johnson 2017) identified some factor groups with potentially significant effects which nevertheless prevented Model 3 from converging. These are shown in Table 4.6.

Factor groups listed as having potentially significant effects and when individually added to Model 3 prevent it from converging			
	***p* values**		
Potentially significant factor groups	**Factor weights**	**% *avoir***	**Total N**
Type of subject	$p = 0.032$		
Ça	0.682	73,9	23
Strong pronoun	0.658	46,4	28
Weak pronoun	0.56	32,6	558
Relative *qui*	0.333	27	37
Common noun DP	0.276	27	37
Polarity of the sentence	$p = 0.037$		
Positive	0.661	34,2	658
Negative	0.339	28	25

Table 4.6 Factor groups that are identified as potentially significant for Model 3 in *Shiny Rbrul* (Johnson 2017) but when they were individually added as current predictors the model could not converge

Table 4.6 shows that the factor groups 'Type of subject' and 'Polarity of the sentence' are likely to be playing a small role in the alternation (both with a significant $p < 0.05$). With the 'Type of subject', *ça* and strong pronouns in dislocation (*moi, toi, nous autres*, etc.) both favour the use of *avoir* with respective FWs of 0.682 and 0.658, as opposed to relative *qui* and common noun DPs which disfavour it, with respective FWs of 0.333 and 0.276. Weak pronouns have almost an equal chance of favouring either auxiliary verbs, with a FW very close to 0.5 (0.56). Since both Model 2 and Model 3 have confirmed the very high level of significance of '(In)animacy of the subject' as a factor group ($p < 0.001$), with inanimate subjects favouring *avoir*, it is therefore unsurprising that inanimate subject *ça* favours *avoir* the most. Even though the factor group 'Type of subject' was not identified as having a significant or potentially significant effect in the comparative model (Model 2), this factor group was tested by Sankoff (2019: 223-224) on the combined 1971, 1984, and 1995 Montréal data, as mentioned in §2.4.2.3.10., and my results (presented in Table 4.6) are again consistent with

hers, i.e. *avoir* was found to be significantly more likely with 3SG subjects *ce/ça* ($p < 0.05$) and relative *qui* also disfavoured *avoir* use ($p < 0.001$) in her data.

As to the factor group 'Polarity of the sentence', a positive sentence again favours *avoir*, with a FW of 0.661, and a negative one disfavours it, with a FW of 0.339. In the previous analysis of other potentially significant factor groups that nevertheless made the comparative model (Model 2) unable to converge, 'Polarity' was also identified by *Shiny Rbrul* (Johnson 2017).

4.2.3. Distributional analysis of the variables tested as random effects

In this section, I present a distributional analysis with descriptive statistics of the two independent variables that were tested as random effects in the multivariate analyses, namely the individual speaker and the lexical item.

4.2.3.1. Individual speaker

The factor group 'Individual speaker' was included as a random effect in the multivariate analyses for reasons explained in §3.2.2.1., §3.5.2., and §4.2.2.1, and the patterns observed in Models 2 and 3 are robust despite important variation across speakers. It is this interspeaker variation that is analysed in detail here.

Among the 48 speakers of the corpus, 46 speakers displayed auxiliary alternation with at least one verb. The remaining two speakers were categorical users of auxiliary *être* and, as was mentioned earlier, no speaker used *avoir* exclusively. Table 4.7 shows the overall rates of *avoir* selection per speaker, in decreasing order. Interspeaker variation is considerable, with total rates ranging from 0% to 45,8% of *avoir* selection.

Speaker pseudonym	N of *avoir* tokens	Total N of tokens	% *avoir* selection
Philippe	11	24	45,8
Linda	17	43	39,5
Yves	8	44	18,2
Sylvain	14	89	15,7
Maxime	5	32	15,6
Sara	11	75	14,7
Joël	7	48	14,6
Sophie	16	117	13,7
Gaëtan	13	97	13,4
Jacynthe	5	40	12,5
Rachel	10	90	11,1
Dominic	11	104	10,6
Steve	7	67	10,4
Jean-François	6	60	10
Alexis	5	54	9,3
Julie	13	140	9,3
Madeleine	9	106	8,5
Nathan	6	72	8,3
Virginie	3	38	7,9
Marc	5	64	7,8
Annie	4	65	6,2
Kim	5	83	6
Jessica	3	52	5,8
Marie-Jeanne	4	70	5,7
Martin	5	87	5,7
Johanne	5	89	5,6
Martine	5	90	5,6
Sabrina	4	73	5,5
Caroline	11	202	5,4
Justin	3	57	5,3
Denise	4	77	5,2
Guylaine	5	98	5,1
Marie-Laurence	1	21	4,8
Annouck	6	129	4,7
Charles-Antoine	3	68	4,4
Mathieu	2	45	4,4
David	4	92	4,3

Florence	2	70	2,9
Hugo	1	36	2,8
Denis	4	152	2,6
Richard	1	42	2,4
Paul	2	113	1,8
Carl	1	58	1,7
Amélie	2	115	1,7
Christine	1	84	1,2
Mario	1	100	1
Francis	0	25	0
Marie-Lou	0	83	0

Table 4.7 Speakers of the Montréal corpus of 2016 ranked in decreasing order of *avoir*-selection rates

In terms of intraspeaker variation, it is much more difficult to detect patterns because in many cases the periphrastic tense tokens appeared only once per speaker and per verb. However, it seems to be the case that most speakers were indeed categorical users of *avoir* for a certain number of verbs during the recording session, as illustrated in Appendix 4B, Table 9.2. The table first ranks the speakers by decreasing order of their *avoir*-selection rate, and additionally by verb (in decreasing order of their overall rate of *avoir* in the corpus) in order to verify whether each speaker's lexical constraints are the same. Table 9.2 shows that out of the 48 participants, nine speakers (Marie-Jeanne, Marie-Laurence, Hugo, Richard, Paul, Carl, Amélie, Christine, Mario) only produced one token of *avoir* and four speakers (Philippe, Joël, Nathan, Denise) used *avoir* categorically for certain verbs and *être* categorically for the others. The table also shows that approximately a quarter of the 48 speakers (Yves, Maxime, Gaëtan, Rachel, Jean-François, Virginie, Martine, Sabrina, Caroline, Guylaine, Annouck, Charles-Antoine, Mathieu) shared the same lexical constraints as the whole speech community: the *avoir*-selection rates per verb of these 13 speakers followed the same *avoir* rates order per verb as is presented below in Tables 4.8 and 4.9. Table 9.2 also shows that the range of verbs that display

auxiliary alternation, or that are categorically conjugated with *avoir*, varies greatly between the 46 speakers that were variable users (from 1 to 8 verbs conjugated with *avoir* per speaker).

Table 4.8 presents a summarized version of Table 9.2 (Appendix 4B) by showing the spread of *avoir* usage across the corpus: it displays the number of speakers with categorical *avoir* use, variable use, and categorical *être* use for each verb.

Lexical verb	**N of *avoir* tokens**	**Total N of tokens**	**% of *avoir***	**N of speakers with categorical *avoir***	**N of speakers with variable *avoir***	**N of speakers with categorical *être***
(re)déménager	57	71	80,3	15	8	3
(re)passer	41	70	58,6	16	6	8
(re)monter	18	46	39,1	7	4	13
(re)tomber	42	132	31,8	8	19	14
(re)descendre	11	37	29,7	4	3	13
(re-)rentrer	26	122	21,3	5	10	20
(res)sortir	18	93	19,4	2	12	20
rester	19	98	19,4	6	6	24
retourner	7	58	12,1	4	3	22
(re)partir	7	259	2,7	0	6	38
arriver	2	308	0,6	0	2	46
(re)venir	2	325	0,6	0	2	44
aller	3	755	0,4	0	3	42[174]

Table 4.8 Number of speakers who used *avoir* categorically during the recordings, or who displayed variable auxiliary selection or categorical *être* use per lexical verb (in decreasing order of overall *avoir* selection per verb)

Consequentially, Table 4.8 shows that the total number of categorical *avoir* users decreases alongside decreasing *avoir* rates per verb, and the total number of categorical *être* users follows the opposite trend. Table 4.8 also shows that the verbs *(re)déménager* and *(re)passer* have the

[174] The remaining three speakers used *avoir été* systematically as an equivalent of *être allé* (see §4.4.1. for a detailed discussion of *avoir été* distribution).

most categorical *avoir* users, and that verbs *(re)tomber*, *(re-)rentrer,* and *(res)sortir* have the most variable auxiliary users.

4.2.3.2. Lexical item

It was not possible to test the effect of the lexical item as a fixed one with a multivariate analysis because no model could converge when the factor group 'Lexical verb' was not coded as a random effect, most likely because it contains too many factor levels (13 individual lexical items). The distribution of *avoir* selection by lexical verb is given in Table 4.9 (and the detailed distribution of *avoir* selection by lexical item showing how each verb and its iterative counterpart behave individually is presented below in Table 4.11).[175]

Lexical verb	**N of *avoir***	**Total N**	**% of *avoir***
déménager / redéménager	57	71	80,3
passer / repasser	41	70	58,6
monter / remonter	18	46	39,1
tomber / retomber	42	132	31,8
descendre / redescendre	11	37	29,7
rentrer / re-rentrer	26	122	21,3
sortir / ressortir	18	93	19,4
rester	19	98	19,4
retourner	7	58	12,1
partir / repartir	7	259	2,7
arriver	2	308	0,6
venir / revenir	2	325	0,6
aller	3	755	0,4
Total	**253**	**2374**	**10,7**

Table 4.9 Distribution of *avoir* selection in percentages per lexical item (in decreasing order of %)

[175] See §4.2.3.2. for a discussion of the role of derivational morphology, specifically the iterative prefix *re-*, in auxiliary selection.

Rates of *avoir* selection vary widely: the percentages span from 0,4% to 80,3%. Similarly, Sankoff & Thibault (1977: 96) reported a wide distribution ranging from 0,7% to 90%. Such spread can also be observed in other varieties of French: 24% to 90% in Vermont (Russo & Roberts 1999: 77), 22% to 91% in Welland-Sudbury and Rayside-Balfour (Canale *et al.* 1978: 516), 8% to 94% in Ottawa-Hull (Willis 2000: 55), and 8% to 78% in Chicoutimi-Jonquière (Renaud & Villeneuve 2008).

Additionally, Table 4.10 shows that the two Montréal French studies as well as studies on other varieties of French have not all recorded the same number of verbs exhibiting auxiliary alternation. As Willis (2000: 54) states, these findings suggest that, contrary to what prescriptive works assert by simply listing *Ê*-verbs, "the choice of auxiliary in French *composé* tenses is not simply determined by the lexical verb used. If that were the case, I would expect the verbs showing variation in the different varieties of French to be identical."

Lexical verb	Montréal (2020)	Montréal (1977/1980;[176] 2019)	Chicoutimi-Jonquière (2008)	Ottawa-Hull (2000)	Vermont (1999)	Welland, Sudbury Rayside-Balfour (1978)
aller	x	x			x	x
apparaître				x		
arriver	x	x	x	x	x	x
commencer				x		
décéder				x		
déménager	x	x		x	x	
demeurer		x		x	x	
descendre	x	x	x	x		
devenir			x	x		
entrer	(x)	x		x		
monter	x	x	x	x	x	
partir	x	x	x	x	x	x
passer	x	x	x	x	x	
redéménager	x					
redescendre			x	x		
remonter	x		x			
rentrer	x	x	x	x	x	x
repartir	x		x	x		
ressortir	x					
rester	x	x	x	x	x	x
retomber			x			
retourner	x	x	x	x	x	
revenir	x	x	x	x		x
sortir	x	x	x	x	x	x
tomber	x	x	x	x	x	x
venir		x	x	x	x	x
Total number of verbs	**17**	**16**	**17**	**21**	**13**	**9**
Total number of tokens	**2341**	**2289**	**487**	**2001**	**453**	**417**

Table 4.10 List of intransitive *Ê*-verbs displaying auxiliary alternation by speech community

[176] Sankoff & Thibault (1977) had also studied *changer* but excluded it from their 1980 paper.

While it is almost certain that a lexical effect is present, it seems obvious from the data in Table 4.10 that it differs across communities. This, however, could be an effect of inevitably small numbers of tokens collected for certain verbs, even in the largest studies.

For the purpose of mixed-effects modeling, the verbs in my study had to be collapsed with their iterative counterpart in order to create a converging model, even when the lexical item was set as a random effect.[177] However, as Willis (2000: 63) outlined in her work on auxiliary alternation in Ottawa-Hull French, there is no way of knowing whether "speakers consider each of these verbs to correspond to a single lexical entry, or to a number of separate lexical items". The detailed distribution of *avoir* selection per lexical item, showing how each verb and its iterative counterpart behave individually, is presented in Table 4.11.

[177] Future work on auxiliary alternation could run separate multivariate analyses for each verb in order to determine whether the same conditioning factors are at play in the data of each individual verb.

Lexical verb[178]	N of *avoir*	Total N	% *avoir*
redéménager	4	4	100,0
déménager	53	67	79,1
passer	41	69	59,4
repasser	0	1	0
remonter	5	9	55,6
monter	13	37	35,1
tomber	42	126	33,3
retomber	0	6	0
descendre	11	34	32,4
redescendre	0	3	0
rentrer	26	120	21,7
re-rentrer	0	2	0
ressortir	1	5	20
sortir	17	88	19,3
rester	19	98	19,4
retourner	7	58	12,1
repartir	2	20	10
partir	5	239	2,1
revenir	2	151	1,3
venir	0	174	0
arriver	2	308	0,6
aller	3	755	0,4

Table 4.11 Distribution of *avoir* selection in % per intransitive lexical item, separating iterative items (in decreasing order of % *avoir*)

Since the iterative tokens were so few in number, it is difficult to determine with accuracy how differently the individual lexical items behave. But by grouping all the morphologically-derived forms together it is possible to determine whether iterativity might be playing a role in the auxiliary alternation.

[178] Unsurprisingly, the corpus had no occurrence (with neither *être* nor *avoir*) of the iterative forms *re-rester, re-retourner, re-arriver*, and *re-aller*.

The distribution of *avoir* selection based on iterativity is presented in Table 4.12.

	N of *avoir*	**Total N**	**% *avoir***
Non-iteratives tokens	240	2173	**11**
Iteratives tokens	14	201	**7**

Table 4.12 Comparison of *avoir*-selection rates for non-iterative and iterative verbs

Since this slight difference in *avoir*-selection rates between non-iterative tokens and iterative ones (11% vs 7%) does not translate in a statistically significant factor group in any of the mixed-effects models built with *Shiny Rbrul* (Johnson 2017), it might either be the case that iterative tokens have the same probability of surfacing with *avoir* as non-iterative ones and/or that speakers do not perceive all iterative forms as following similar auxiliation pattern.

4.2.4. Distributional analysis of variables tested as fixed effects

In this section, I provide the descriptive statistics for two factor groups that were tested as fixed effects in the multivariate analyses with *Shiny Rbrul* (Johnson 2017) presented in §4.2.2.2. and §4.2.2.3. and were not found to have a significant effect on the data. The two independent variables that are examined in more detail are the effects of the contact with English and of the person and number of the subject/auxiliary verb. These two variables are further surveyed here in order to compare Montréal French with auxiliation patterns observed in other Romance languages, in which there is no close contact with English and where the person of the subject is often cited as a parameter of variation as outlined in §2.2.1.1.3., and to illustrate why they were not selected as having a statistically significant effect in the multivariate analyses.

4.2.4.1. Contact with English

It was postulated in §3.2.2.6. that since most studies on auxiliary alternation in French had been carried out in North America and in primarily English-speaking communities, since spoken Québécois French is so readily associated with an important use of Anglicisms (mostly semantic, lexical, and syntactic ones), and since HAVE is the sole auxiliary verb of the present perfect tense in English, it was probable that contact with English had an effect on auxiliary alternation. While this factor group ('Bilinguals vs Non-bilinguals') was statistically tested in Models 2 and 3 in §4.2.2. and was not selected as having a significant effect on *avoir* usage, this section will show the various distributions of *avoir* rates by English proficiency level in order to explain what lies behind the statistical findings. As outlined in §3.2.2.6., since Poplack (1997) states that one of the ways to reveal the effect of contact with English is to observe the behaviour of the speakers who have the best knowledge of English, the present study examines the reported frequency of use of English in the daily life of the speakers combined with the type of bilingualism involved.

Having established that the LOW SPS is the only social predictor for *avoir* usage, Table 4.13 shows the rates of English proficiency from highest to lowest with a distribution of the SPS subgroups. It is worth noting that the 'Non-bilingual' label might still encompass various levels of lower English proficiencies.

Levels of English proficiency	**N of speakers**	**Average % *avoir* of the speakers**
Native bilingualism	6 (2HIGH, 3MID, 1LOW)	12,2 (range: 5,5-14,7)
Acquired bilingualism	14 (6H, 5M, 3L)	13 (1-15,6)
Non-bilingual	28 (8H, 8M, 12L)	11,2 (0-45,8)

Table 4.13 Distribution of *avoir* selection based on three levels of English proficiency

Table 4.13 shows that there is only a very small difference in *avoir* rates, between 1% and 2%, between the three different levels of English proficiency, even though the distribution of LOW SPS speakers is uneven, and there is a lot of variation around the mean, especially in the 'Non-bilingual' category (see ranges given in Table 4.13).

After grouping together the native and acquired bilingualism speakers, the distribution of LOW SPS speakers remains uneven: there is an equal number of HIGH and MID SPS speakers (8 in each subgroup), as well as 4 LOW SPS speakers, but there are 12 LOW SPS in the non-bilingual category. Table 4.14 shows the average rates of *avoir* per grouping.

Merged levels of English proficiency	**N of speakers**	**Average % *avoir* of the speakers**
Bilingual	20 (8H, 8M, 4L)	12,6 (1-15,6)
Non-bilingual	28 (8H, 8M, 12L)	11,2 (0-45,8)

Table 4.14 Distribution of *avoir* selection based on English bilingualism (yes/no)

Table 4.14 suggests that on average bilinguals and non-bilinguals have a very similar rate of *avoir* selection even though the LOW SPS speakers are three times more numerous in the 'Non-bilingual' category. Moreover, a lot of variation around the mean was recorded (see ranges given in Table 4.14). It is therefore not surprising that when these collapsed groupings were tested statistically, as reported in §4.2.2., 'Bilingualism' was not selected by *Shiny Rbrul* (Johnson 2017) as a factor group with a significant effect ($p > 0.05$ in Models 2 and 3). It can therefore be established that contact with English does not play a role in the recorded auxiliary alternation.

4.2.4.2. Person and number of the subject/auxiliary

While the factor groups 'Person and number of the subject' and 'Form of the auxiliary' were not selected by *Shiny Rbrul* (Johnson 2017) as factor groups with significant effects[179] in neither Model 2 or 3 (and when included prevented any model from converging), the inanimacy of the subject was the best predictor for *avoir* selection in Model 2 since the factor group '(In)animacy of the subject' had the smallest *p* value (p = 4.97e-07). The only persons that can have inanimate referents are third-persons (both singular and plural), and because of the presence of this obvious confound, it is therefore not unexpected that the effect of this factor group was not significant. The distribution of *avoir*-selection rates per person is presented in Table 4.15.

Person and number of the subject	**Total N of *avoir***	**Total N**	**% *avoir***
1SG	98	942	10,4
2SG	7	53	13,2
3SG	76	615	12,4
1PL '*on*' (form of the aux. in 3SG)	58	622	9,3
2PL	0	5	0
3PL	9	109	8,3

Table 4.15 Distribution of *avoir* selection by person and number of the verb (in decreasing order)

Table 4.15 shows that the third-person results do not stand out as very different from the other persons, and so, if anything, the finding that (in)animacy plays a major role in the alternation is surprising in the light of the confound. Table 4.15 also shows that the difference in *avoir* rates between the various persons of the verb does not appear substantial at first

[179] Future studies on auxiliary alternation could possibly group persons in various ways: 1st vs 2nd vs 3rd, 1st and 2nd vs 3rd, 1st vs all others, 1SG vs all others, etc.

glance, even though it has been shown in Table 4.6 that subject *ça* (3SG) is likely to favour *avoir* selection.

Moreover, the presence of substantial interspeaker and intraspeaker variation makes impossible the creation of one-form-per-cell conjugation paradigms for verbs, such as those compiled by Manzini & Savoia (2005) or by Štichauer (2018) for person-based perfective auxiliation systems in some Italo-Romance dialects. Given my findings, there is scope to question how sociolinguistically realistic the kind of data used by these linguists really is.

4.2.5. Distributional analysis of the independent variables not tested in the multivariate analyses

I present here distributional analyses with descriptive statistics of the social and linguistic independent variables not tested in the multivariate analyses, in order to establish whether some patterns emerge in terms of auxiliary choice. These variables were not tested in the multivariate analyses because their analysis is better suited to descriptive statistics, as explained in §3.2.2. and §3.2.3.

4.2.5.1. Sociolinguistic independent variables

The following social variables were not included in the multivariate analysis because after proceeding by judgement sampling to obtain 48 speakers in 12 balanced cells, the data obtained for these variables were subsequently not balanced. The influence of social variables comprising stylistic variation, contact with other varieties of French, and contact with other languages is examined in detail in this section.

4.2.5.1.1. Stylistic variation

As detailed in §3.2.2.7., in order to create an index scale of stylistic variation with five levels, the speakers were first divided into groups based on whether or not they had been interviewed in a pair (the pairs consisted of close friends, life partners, siblings, etc.), and they were then divided by three levels of familiarity with the interviewer.

Figure 4.1 shows the average rates of *avoir* selection per level of formality (a total score of 1, 2, 3, 4, or 6), ranked by decreasing level of formality. The red line was added by hand to approximate the expected pattern based on the hypothesis that when the level of formality decreases, the average *avoir*-selection rates will increase.

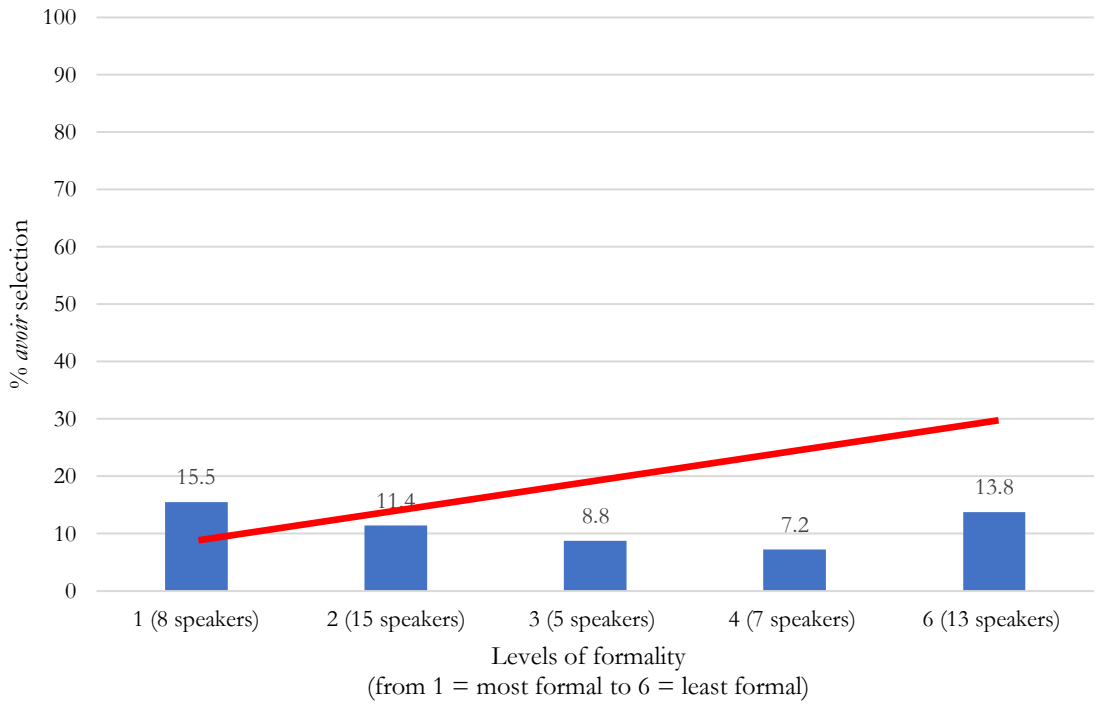

Figure 4.1 Percentage of *avoir* selection based on the level of formality (in decreasing order of formality) contrasted with the expected pattern

Figure 4.1 actually shows that no linear pattern can be observed. Between levels 1 and 4, it is actually the opposite trend that emerges. The average *avoir*-selection rate of the speakers in level 6 (those who presumably felt the most comfortable being interviewed by me) do not even surpass the average rate of the speakers with the most formal level. Since the lowest SPS is the strongest predictor of *avoir,* one could have assumed that the observed pattern could have simply been due to how my own personal contacts were unevenly distributed in terms of SPS, i.e. that there would be more higher SPS speakers in the closest levels of familiarity. However, the cells were fairly balanced in terms of SPS, as Table 4.16 shows:

Level of familiarity	**N of LOW speakers**	**N of MID speakers**	**N of HIGH speakers**	**Total**
No personal connection to the interviewer	8	5	6	19
Participant knew the interviewer through a close contact	3	4	4	11
Participant knew the interviewer on a personal level	5	7	6	18

Table 4.16 Distribution of speakers based on the SPS and their level of familiarity with the interviewer

Table 4.17 shows the distribution of speakers based on the SPS and whether or not they were interviewed in a pair.

Interview situation	**N of LOW speakers**	**N of MID speakers**	**N of HIGH speakers**	**Total**
Interviewed individually	7	6	4	17
Interviewed in a pair	9	10	12	31

Table 4.17 Distribution of speakers based on the SPS and whether or not they were interviewed in a pair

Table 4.17 shows that there were almost twice as many speakers who were interviewed in a pair (31) than of those who were not (17), and that there were four times more HIGH SPS speakers who were interviewed in a pair than those who were interviewed individually, which may have counterbalanced the informality of their speech.

4.2.5.1.2. Contact with other varieties of French

As mentioned in §3.2.2.6., the 48 speakers were asked in the personal information questionnaire whether they had had contact, over an extended period of time in their life, with speakers of other varieties of French ("from France, Belgium, Switzerland, Haiti, Senegal, Benin, etc.", see Appendix 3B). Only the contacts that included a life-partner, a parent, a long-time colleague, and a very close friend were considered. It was hypothesized that if speakers were exposed to varieties of French that do not display auxiliary alternation as predominantly as Canadian French, they would produce lower rates of *avoir* selection. Table 4.18 shows the distribution of *avoir* selection based on whether the speakers had prolonged contact with native speakers of other varieties of French.

Contact with other varieties of French	**N of speakers**	**Average % *avoir* of the speakers**
No	27 (6H, 10M, 11L)	15,2 (1,7-45,8)
Yes	21 (10H, 6M, 5L)	7,5 (0-13,7)

Table 4.18 Distribution of *avoir* selection based on contact with other French varieties

Table 4.18 shows that speakers who have not been in close contact with other varieties of French produce an average *avoir* rate that is twice that of the speakers who had been exposed to other varieties of French. However, this seemingly stark effect can most probably be

explained by an interaction with the SPS: there were more than twice as many LOW SPS speakers in the 'no contact' category than in the other one, and almost twice as many HIGH SPS speakers in the 'contact' category. In my corpus, it is therefore unlikely that the contact with other varieties of French had any effect on *avoir* selection because of the partial confound with the SPS, but also because a lot of variation around the mean was recorded (see ranges given in Table 4.18).

4.2.5.1.3. Contact with other languages

It was not possible to properly test whether the contact with other languages had an impact on auxiliary alternation because only four speakers reported speaking another language (apart from English): one speaker was a native speaker of both Cantonese and Vietnamese, and three reported speaking Spanish, including one as a native speaker (Sabrina).

Given the lack of BE-auxiliary in Spanish, it could have been the case that this native bilingual speaker (French and Spanish) exhibited a higher rate of *avoir* selection, but her data yielded a fairly low rate of 5,5%. This can probably be explained by the fact that she is a HIGH SPS speaker.

4.2.5.2. Linguistic independent variables

The following sections of this chapter will cover various linguistic variables that have not been included in the multivariate analysis, either because their factor levels comprised too few tokens to be statistically tested (e.g. verb meanings) or because their analysis was more suited to a descriptive approach (e.g. frequency, possibility of parallel adjectival use). They will be examined in detail, with descriptive statistics, since they have proven to be factors of influence in other studies of auxiliary alternation (see §2.4.2.3.).

4.2.5.2.1. Verb meanings

Out of the 13 verbs studied, some can have both a 'core' meaning, usually motional or relating to a change of state, and a figurative meaning, as explained in §3.2.3.2. Four verbs had enough *avoir* tokens in the corpus for me to be able to verify whether different meanings could explain some of the auxiliary alternation recorded in 2016.[180] The verbs which are examined here are *sortir*, *tomber, passer*, and *rentrer*. Illustrative examples of all uses were presented in §3.2.3.2. The distribution of *avoir* selection per meaning of *sortir* is presented in Table 4.19.

Verb	**N of *avoir* tokens**	**Total N**	**% *avoir* selection**
(Res)sortir avec quelqu'un 'to date, to go out with someone'	8	15	53,3
(Res)sortir (core meaning)	10	75	13,3

Table 4.19 Distribution of *avoir* selection for the verb *sortir*

Table 4.19 shows that more than half of the tokens of *sortir* when used to mean 'to date, to go out with' surfaced with *avoir,* as opposed to 13,3% of those with the 'core' motional meaning of *sortir*.

As to the verb *tomber*, for the purpose of this analysis, all figurative uses have been combined and counted together. The distribution of *avoir* selection per use of *tomber* is presented in Table 4.20.

[180] Future studies on auxiliary alternation could regroup all the 'non-core' tokens of intransitive *Ê*-verbs and test that statistically, against all 'core' uses, as an independent variable in *Shiny Rbrul* (Johnson 2017).

Verb	N of *avoir* tokens	Total N	% *avoir* selection
(Re)tomber (figurative use)	10	42	23,8
(Re)tomber (core meaning)	31	83	37,3

Table 4.20 Distribution of *avoir* selection for the verb *tomber*

In contrast with the verb *sortir,* the figurative uses of *tomber* (23,8%) seem to select *avoir* in a smaller proportion than its core meaning (37,3%).

With *passer,* the idiomatic expression *passer proche (de mourir)* 'to come close to dying' has been used five times in total but all its tokens surfaced with *avoir*. The distribution of *avoir* selection per use of *passer* is presented in Table 4.21.

Verb	N of *avoir* tokens	Total N	% *avoir* selection
Passer proche (de mourir) (lexicalized meaning)	5	5	100
Passer (core meaning)	36	64	56,3

Table 4.21 Distribution of *avoir* selection for the verb *passer*

The most dramatic difference in *avoir*-selection rates is seen in the various usages of the verb *rentrer*. While the analysis below does not oppose a 'typical' figurative use of *rentrer* to its 'core' meaning, it rather examines the lexical field in which the verb is used. The verb *rentrer* was used with *avoir* in almost every instance where a speaker is discussing a bicycle crash or a car crash.

Most of the 'accident' tokens of *rentrer* were uttered when the speakers were asked my version of the Danger of Death question (see §3.1.3.). Table 4.22 summarizes the distribution of *avoir*-selection rates depending on the usage of the verb *rentrer*.

Verb	N of *avoir* tokens	Total N	% *avoir* selection
Rentrer (de)dans (accident)	17	19	89,4
Rentrer (other meanings)	9	101	8,9

Table 4.22 Distribution of *avoir* selection for the verb *rentrer*

Table 4.22 shows that *avoir* surfaces almost categorically (89,4%) when speakers use the verb *rentrer* to discuss a crash of some kind. Interestingly, out of these 17 tokens, eight had an object pronoun surface before the auxiliary verb, as shown in (66).

(66) *Faque c'était juste comme un effet domino qui a fait que moi je* ***leur ai rentré*** *dedans.* (Annie)
'So it was just like a domino effect which made me run into them [with a car]'

It should also be noted that the factor 'Pronoun before the auxiliary' was identified above as a highly significant predictor of *avoir* selection (see Tables 4.2 and 4.5), and no tokens of the other meanings of *rentrer* had any pronoun appearing before the auxiliary.

From this preliminary analysis,[181] it is obvious that verb semantics are at play in the auxiliary alternation recorded with *passer, sortir* and *rentrer*. It is also worth pointing out that it is not a coincidence perhaps that the last two verbs pattern similarly, given that the pair are semantically opposed. As to *tomber,* it is interesting to note that it is rather its core meaning that shows a higher *avoir* rate.

[181] Future studies on auxiliary alternation could code each token depending on the verb meaning, i.e. whether it refers to a state, change of state, or a motion, as Sankoff (2019) has done when she revisited her 1971-1984-1995 data. This was found by Sankoff (2019) to be a significant factor group in that combined corpora data.

4.2.5.2.2. Frequency of use

As mentioned in §3.2.3.3., the frequency of use of the various verbs is studied here by using their overall corpus frequency, following Sankoff & Thibault (1977), and it was expected that high frequency verbs would select *avoir* in lower proportions than low-frequency ones. Given that iterative tokens will surface in lower quantities than non-iterative ones, it was decided that the potential influence of the frequency of use would be examined by analysing iterative and non-iterative lexical items separately. Otherwise, the frequency effect would be difficult to disentangle if both forms were not distinguished from each other. Figure 4.2 shows the overall corpus frequency of the different lexical verbs studied, superimposed with their respective rates of *avoir* selection.

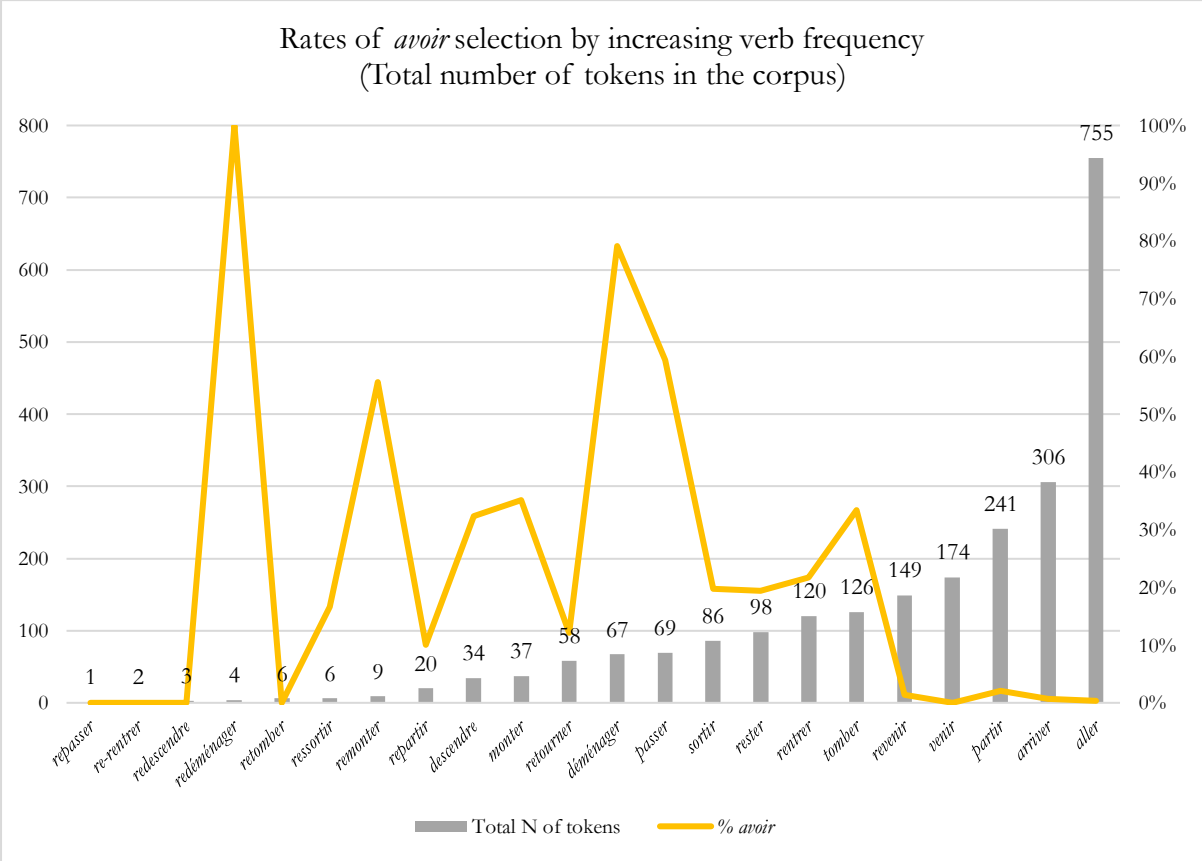

Figure 4.2 Percentage rates of *avoir* selection by verb corpus frequency

What can be inferred from Figure 4.2 is that frequency is not directly correlated with *avoir* selection, at least not with low-frequency verbs (listed to the left-hand side of the figure) where there is a lot of variability. If the correlation were direct, I would expect to see a negative non-linear yellow line, i.e. a negative slope. Frequency does not appear to be a good indicator of auxiliary choice overall, except for high frequency verbs, which have very low rates of *avoir* selection. A clear negative slope can however be seen from the middle of the graph, starting with *déménager*[182] and *passer*, that rises slightly with verbs *rentrer, tomber,* and *partir* before almost reaching 0% to the right extremity of the figure. This result confirms findings by Sankoff & Thibault (1980: 334-335), who had noted that very frequent verbs, such as *aller*, *venir,* and *arriver,* very rarely surfaced with *avoir* in Montréal French in 1971. To this list, I must now add *revenir* as well as *partir*.

4.2.5.2.3. Possibility of parallel adjectival use

The summary of previous findings regarding the possibility of parallel adjectival use presented above in Table 3.7 showed that the literature is inconsistent in the treatment of this independent variable. The verbs *aller* and *venir* are the only two that do not allow adjectival use in this list, and are also the ones that yielded the smallest number of *avoir* tokens in my corpus: *aller* (0,4% of *avoir* selection: 3 out of 755 tokens) and *venir* (0%: 0 out of 174). However, a verb such as *(re)partir* which allows adjectival use according to all studies, or *arriver,* which allows adjectival use according to Canale *et al.* (1977) and to my own judgement, almost never surfaced with *avoir* in the 2016 data (2,7% of tokens for *(re)partir* and 0,6% for *arriver*). So even though this factor group was the best predictor of the auxiliary alternation observed

[182] It is likely that the verb *déménager* surfaces disproportionately frequently in this corpus because, in order to ensure the collection of a sufficient quantity of tokens of motion verbs in the past, specific questions relating to house moves were asked to all participants.

in Ottawa-Hull (Willis 2000) as well as the best linguistic predictor for Sankoff & Thibault's data, it is apparent that this is definitely not the case with the 2016 Montréal data.

4.2.6. Comparison with the Sankoff-Cedergren Montréal corpus (1971), the Montréal 1984 corpus, and the Montréal 1995 corpus

Because of the way in which my sociolinguistic interviews were structured in terms of targetted conversation topics, my sample comprises more tokens (2374 tokens) than were collected by Sankoff & Thibault in 1971 (2193 tokens), even though their corpus included two-and-a-half times more speakers than mine (120 vs 48 speakers). Nevertheless, a real-time comparison of the speech community at two points in time shows that the rate of *avoir* selection is in 2016 only about a third of what was recorded in 1971 (from 32,8% to 10,7%), as mentioned in §4.2.1.

Sankoff & Thibault (1977) recorded auxiliary alternation in two verbs that, in contrasting ways, are categorical in their auxiliary selection in 2016: *demeurer* and *venir*. The verb *demeurer* has not been included in my study because its only two tokens were conjugated with *avoir,* while *venir* was categorically used with *être* (174/174). It therefore appears that the change (potentially back) to *être* with *venir* has reached completion, and *venir* now behaves like *devenir, naître,* and *mourir,* which displayed no auxiliary alternation whatsoever either in 1971 (Sankoff 2019: 203) or in 2016. Interestingly, in 1977 Sankoff & Thibault had recorded (Sankoff 2019: 205) high rates of *avoir* selection in some collocations involving *venir*, such as *venir à bout* 'to succeed' (42% of 31 tokens) and *venir au monde* 'to be born' (13% of 119 tokens), however none of these idiomatic expressions were recorded in the 2016 corpus.

Figure 4.3 illustrates the distribution of *avoir* selection for the various lexical verbs studied in 1971 and in 2016. The verbs are listed in increasing order of *avoir* selection as recorded in 1971.

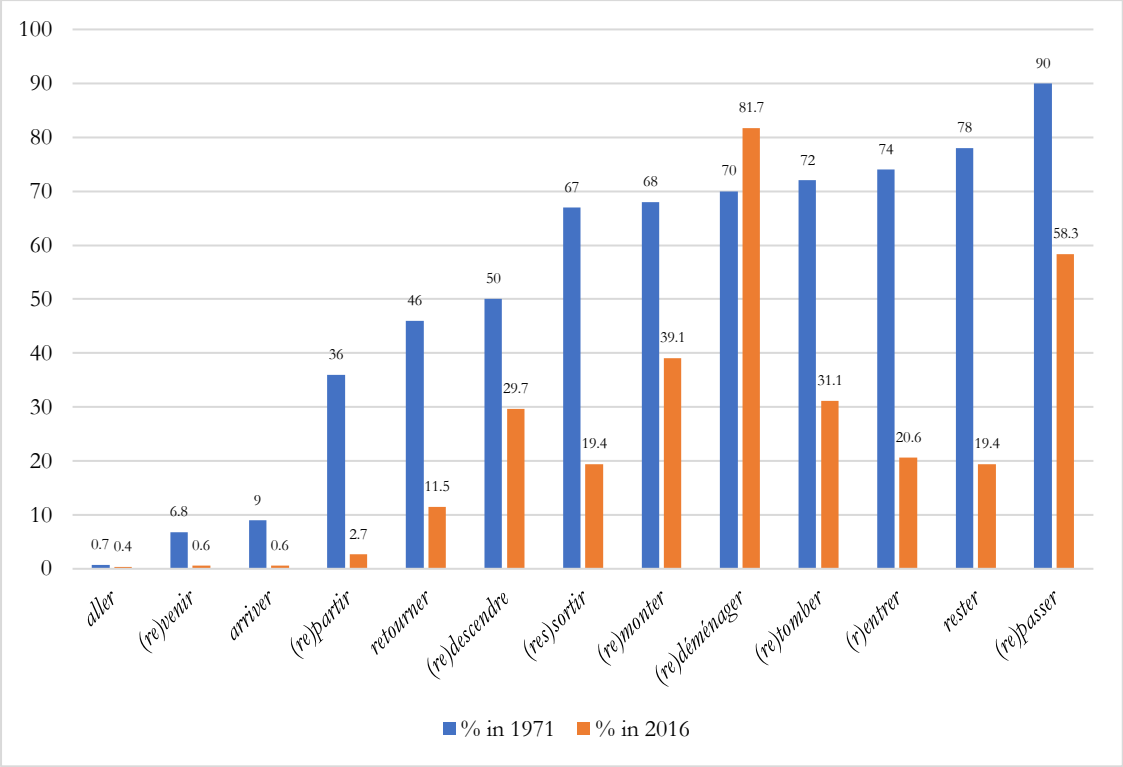

Figure 4.3 Distribution of *avoir* selection (in %) per lexical verb in 1971 and 2016

It is worth noting that the verb *(re)déménager* is the only one that shows an increase in *avoir* use over time, but that *(re)déménager* has always taken *avoir* in European varieties of French and standard French. This might be an indication that speakers who use auxiliary *être* with *(re)déménager* are in fact hypercorrecting. Otherwise, all the other verbs show a dramatic drop in their selection of *avoir*.

The statistical analysis performed by Sankoff & Thibault (1977) on the 1971 data revealed some effect of the frequency of use, the level of education, the socioeconomic class as well as a small gender effect (see §2.4.2.2.1.). Moreover, while Sankoff & Thibault (1977) had posited the impact of the possibility of transitive use and of adjectival use with copula, it is unclear whether these effects were tested statistically on their data. My findings confirm

their results in part, as the frequency of use of the verb (only with very frequent verbs) and the SPS (which included a measure of the level of education) did appear to play a role in the alternation. However, in my data, only one significant sociolinguistic factor group (the SPS) was identified by *Shiny Rbrul* (Johnson 2017). As mentioned in §3.5.2., it is likely that this discrepancy is caused by the different way in which statistical software used prior to *R* (*R* Core Team 2004-2019) and *Rbrul* (Johnson 2009) analysed data: they tended to exaggerate the impact of social factors, such as gender and age (Johnson 2009: 363, 376-377). It could be the case that this reduction in the number of significant social variables also has to do with the overall reduction in *avoir* use as usage moves (back?) towards what is now the standard and also to do with the demographic changes in Montréal outlined in §3.1.1.

With Sankoff's updated 2019 study on the three previous corpora (1971, 1984, and 1995), the results of her regression analysis confirmed a change in progress towards the diminishing selection of *avoir* in apparent time. The variable 'Date of birth' of the speakers yielded a significant $p < 0.05$ (Sankoff 2019: 207), and she proved that an 'age grading' interpretation of this effect (i.e. speakers would have reduced their rates of *avoir* selection as they aged) would be incorrect because auxiliary alternation was a case of "lifespan stability in the face of linguistic community change" (Sankoff 2019: 211). Interestingly, the age of the speakers was not a significant factor of influence on auxiliary alternation in the 2016 data, potentially indicating that this change towards *être* is in its final stages and that a small stable variation might persist, presumably only in the speech of LOW SPS speakers. As to the other social variables, the SPS and the level of education continued to emerge as significant in the three previous corpora, both registering a $p < 0.001$ (Sankoff 2019: 207): the speakers that were the most likely to select *avoir* were the least educated and the ones who were chronically unemployed or on public assistance, as revealed in my 2016 data (see Table 4.2 above). As to the gender of the speakers, it did not emerge as significant in the three previous corpora, in a sample equally balanced between men and women (Sankoff 2019: 207), in contrast with the

1977 study; thus reinforcing the chance that this discrepancy was caused by the use of a different statistical software.

Among the linguistic factors tested, Sankoff (2019: 207-208) observed that in the three previous corpora *avoir* was more likely to be selected when the subject was inanimate ($p < 0.01$), when there was intervening material (usually adverbial) between auxiliary and past participle ($p < 0.01$), when the tenses were others than the *passé composé* ($p < 0.05$). These significant effects were also found in the 2016 data (respectively, $p = 4.97e\text{-}07$, $p = 0.00859$, and $p = 0.0115$), with the exception of the *plus-que-parfait* which disfavoured *avoir* in my data. Sankoff (2019: 208) also found that *avoir* was significantly more likely with 3SG subjects *ce/ça* ($p < 0.05$), but was disfavoured with *il/elle* ($p < 0.001$), *on* ($p < 0.01$), and relative subject pronoun *qui* ($p < 0.05$). In my 2016 data, the type of subject was also selected by *Shiny Rbrul* (Johnson 2017) as a factor group with a potentially significant effect ($p = 0.032$) when added to Model 3 (Table 4.6 in §4.2.2.3. above): *ça* also favoured *avoir* and *qui* also disfavoured *avoir,* but weak pronouns did not have an effect on the distribution of the data.

Mixed-effects Model 2 (Table 4.2 in §4.2.2.2.) also identified the significant effects of three additional factor groups that were not tested by Sankoff (2019): whether there was a clitic pronoun between the subject and the auxiliary verb ($p = 0.00166$), whether the action was hodiernal or pre-hodiernal ($p = 0.00692$), whether the verb could be used transitively ($p = 0.0118$). The presence of a clitic pronoun, pre-hodiernal actions, and the possibility of transitive use all favoured *avoir* use. Given the similarities in the recorded effects on the combined 1971, 1984, and 1995 data and on my 2016 data, going back to test whether these three variables have an significant impact on auxiliary selection in the previous three Montréal French corpora would likely reveal that they do.

The striking fact that almost all the same significant linguistic effects have been identified in the two datasets (with 45 years separating the oldest data from the newest) also

suggests that variation in auxiliary selection will also persist in the linguistic contexts that favour *avoir*.

4.3. Pronominal verbs

This section will examine the current state of auxiliary alternation in pronominal verbs in the Montréal French 2016 corpus. As previously stated in §3.4., all 3PL subjects with pronominal verbs were excluded from the analysis because it was very difficult to distinguish phonetically between tokens *s'ont* with *avoir* and *se sont* with *être*. However, in the 3PL subjects that surfaced in *plus-que-parfait*, where the potential homophony cannot occur, there was one occurrence of a pronominal verb (out of five) conjugated with *avoir*, shown in (67).

(67) *Là, y* ***s'avaient arrangé*** *avec, entre eux autres.* (Philippe)
'Then, they had come to an agreement with, among themselves'

The analysis of the 2016 pronominal data will then be followed by a preliminary comparative study of auxiliary alternation in pronominal verbs in the three Montréal corpora (1971, 1984, 1995) described in §2.4.2.1. This preliminary comparative study was conducted with the help of Gillian Sankoff.

4.3.1. Overall variant distribution

The pronominal verbs exhibit an average *avoir*-selection rate of 1,3%. The distribution is shown in Table 4.23, contrasted with the rate observed in intransitive verbs.

Montréal French data (2016)	**N of *avoir* tokens**	**Total N**	**% *avoir***
Pronominal verbs	**17**	**1263**	**1,3**
Intransitive verbs	253	2374	10,7

Table 4.23 Distribution of *avoir* selection in pronominal verbs contrasted with *avoir* selection in intransitive verbs

Even though the total number of pronominal verb tokens recorded is approximately half of the total intransitive verb ones, the former select *avoir* around ten times less often than the latter. Of the 48 speakers, only six conjugated a pronominal verb with *avoir*, and all six showed intraspeaker auxiliary alternation. Table 4.24 shows their respective *avoir*-selection rates for pronominal verbs, alongside their *avoir*-selection rates for intransitive *Ê*-verbs.

Speaker pseudonym	**N of *avoir* tokens with pronominal verbs**	**Total N of pronominal tokens**	**% *avoir* with pronominal verbs**	**% *avoir* with intransitive verbs**
Linda	8	21	**38,1**	39,5
Philippe	3	8	**37,5**	45,8
Jacynthe	2	8	**25**	12,5
Julie	2	49	**4,1**	9,3
Denise	1	27	**3,7**	5,2
Gaëtan	1	42	**2,4**	13,4
Total	**17**	**155**	**11**	**18,2 (46/253)**

Table 4.24 Distribution of *avoir* selection in pronominal verbs and intransitive verbs for the six variable speakers

The *avoir* rates with pronominal verbs have a wide range (from 2,4% to 38,1%) and all speakers, except Jacynthe, have a lower *avoir* rate with pronominals than with intransitive verbs.

In total, 226 different pronominal verbs surfaced in periphrastic tenses in the corpus, of which 110 were hapax legomena. All of them can be found in Appendix 4C, along with their translations and their respective *avoir*-selection rate. Non-standard uses have also been indicated: a single asterisk (*) indicates that the verb is a Québecism and a double asterisk (**) indicates that the verb is also an Anglicism used commonly in Québec. In the cases where some verbs have more than one translation, the translation chosen is the one that best illustrates the use of the verb as it appeared in the corpus.

4.3.2. Distributional analysis of auxiliary *avoir* by independent variable

Since all pronominal *avoir* tokens surfaced with 1SG subjects, with the exception of the token presented in (67), all other persons were removed from the analysis. This initially brought down the total number of pronominal verb tokens analysed from 1263 to 1231. The 42 non-variable speakers were also removed from the analysis, leaving only 63 tokens in total. Hodiernal tokens (taking place within 24 hours of the moment of speech) and tokens which described an action that had taken place within 48 hours (6 tokens) were recoded and collapsed into a single category.

However, with these exclusions, it was not possible to perform a multivariate analysis with the pronominal verb data because many factor levels within factor groups contained too few tokens to be statistically analysed with *Shiny Rbrul* (Johnson 2017). A distributional analysis of the 63 tokens is presented in Table 4.25.[183]

[183] It should be noted that many cells contain very few tokens and that these results are therefore speculative.

Independent variables	N of *avoir*	Total N	% *avoir*
Speaker pseudonym			
Linda	8	9	88,9
Jacynthe	2	5	40
Philippe	2	5	40
Julie	2	17	11,8
Denise	1	11	9,1
Gaëtan	1	16	6,3
SPS			
Low	14	36	38,9
Mid	1	11	9,1
High	1	16	6,3
Age category			
Older (36-65)	14	46	30,4
Younger (18-35)	2	17	11,8
Gender of the speaker			
Female	13	42	31
Male	3	21	14,3
Pronoun before auxiliary			
Pronoun before aux	5	5	100
No pronoun before aux	11	58	19
Intervening element between auxiliary and past participle			
Intervening element	3	8	37,5
No intervening element	13	55	23,6
Tense			
Conditionnel passé	3	3	100
Plus-que-parfait	3	7	42,9
Passé composé	10	53	18,9
Transitive use			
Transitive use allowed	16	60	26,7
Transitive use not allowed	0	3	0
Polarity of the sentence			
Negative	2	5	40
Positive	14	58	24,1

Type of subject			
Strong pronoun	3	6	50
Pro drop	2	8	25
Weak pronoun	11	49	22,4
Clause type			
Embedded	5	15	33,3
Main	11	48	22,9
Hodiernal action			
Pre-hodiernal	11	41	26,8
Hodiernal	5	22	22,7

Table 4.25 Detailed statistical distribution in percentages of the factor groups with pronominal verbs in the 2016 Montréal data (*avoir* rates with 1SG subjects only)

Auxiliary alternation with pronominal verbs is very restricted: the person of the subject is the best predictor of *avoir* selection, as 100% of the *avoir* pronominal tokens analysed surfaced with 1SG subjects, all animate subjects by definition. This finding is in contrast with what was found for intransitive verbs, which were more likely to surface with auxiliary *avoir* when their subject was inanimate. In descriptive terms, the variation is also socially conditioned: among the six speakers that conjugated a pronominal verb with *avoir,* four belonged to the lowest socioprofessional status (38,9%), one to the middle one (Denise, 9,1%), and one to the highest one (Gaëtan, 6,3%). Additionally, only one speaker out of six belonged to the 'Young' category (18-35 years old) and older speakers selected *avoir* three times more frequently (30,4%) than the younger one (Julie, 11,8%). As to gender, female speakers (31%) produced twice the number of *avoir* tokens than their male counterparts (14,3%). In the data used to create mixed-effects Model 2 with intransitive verbs, the reverse gender trend was observed: men produced 12% of *avoir,* whereas women produced 9,9%. With age, no effect was found with intransitive verbs since 'young' speakers and 'old' speakers produced the same *avoir* rates: 10,9% and 10,7%, respectively.

As to linguistic conditioning, the presence of a clitic object pronoun before the auxiliary seems to favour *avoir* (100%) as opposed to when no clitic pronoun appears (19%). The tokens that had an intervening element between the auxiliary and the past participle have an *avoir* rate of 37,5%, whereas those not doing so surfaced with *avoir* with a lower rate (23,6%). As for the influence of tense, tokens in *conditionnel passé* selected *avoir* with a 100% rate, the *plus-que-parfait* ones with a 42,9% rate, and the *passé composé* ones with a 18,9% rate. Hodiernal tokens (22,7%) surfaced with *avoir* slightly less frequently than pre-hodiernal ones (26,8%). All of these linguistic effects were also observed in the case of intransitive verbs (see §4.2.2.2. and §4.2.2.3. above), and even with this much reduced number of tokens and speakers, the effects observed seem robust, bolstering the findings for the intransitive verbs.

With regard to other linguistic factors, with the possibility of transitive use, the tokens which allowed parallel transitive used surfaced with *avoir* with a rate of 26,7% and none which did not allow transitive use selected *avoir*. In terms of polarity, negative sentences were conjugated with *avoir* more frequently (40%) than positive ones (24,1%). The distribution of types of subject reveals that strong pronouns (50%) surfaced with *avoir* more often than pro-drop subjects (25%) and weak pronoun ones (22,4%). Lastly, as for the effect of the clause type, embedded clauses (33,3%) selected *avoir* a bit more often than main clauses (22,9%).

More data could have allowed me to determine whether an alternation varying along the different types of pronominal verbs is observable: French grammars traditionally split pronominal verbs into four semantic categories: *réfléchi* 'reflexive' (e.g. *se laver soi-même* 'to wash oneself'), *réciproque* 'reciprocal' (e.g. *se parler l'un à l'autre* 'to talk to one another'), *passif* 'passive' (e.g. *tous les livres se sont vendus* 'all the books have sold (themselves)'), or *essentiellement pronominal/sans fonction logique* 'lexical reflexives/non-compositional reflexives' (e.g. *s'enfuir* 'to run away': *enfuir* can only exist in the pronominal form) which all correlate with various types of agreement.

4.3.3. Comparison with the Sankoff-Cedergren Montréal corpus (1971), the Montréal 1984 corpus, and the Montréal 1995 corpus

As stated in §1.1., pronominal verbs were excluded from Sankoff & Thibault's 1977 study and were also not included in Sankoff's 2019 update, but in 2016, with the assistance of Gillian Sankoff, it was possible for me to extract the occurrences of *je m'ai* that had surfaced in the three older Montréal French corpora for comparison, as well as all the *je me suis* tokens produced by the variable speakers.

A few methodological limitations should however be mentioned. The access to the three corpora being quite limited, it was unfortunately neither possible to obtain the total number of tokens of *je me suis* that surfaced in the entire corpora (including those produced by non-variable speakers) nor possible to get hold of the social metadata (gender, age, and SPS) of the speakers that used *avoir* with pronominal verbs. It was therefore impossible to obtain a table equivalent to Table 4.23 presenting the overall *avoir*-selection rate with pronominals for the entire corpora, or to test whether my statistically significant factor groups might apply to this dataset as well. Moreover, since the search of the three corpora was only performed for the phrase *je m'ai,* it would be in any case impossible to test whether some likely significant factor groups for the 2016 Montréal French data – whether there is a pronoun before the auxiliary, the tense of the verb, whether there is intervening material between the auxiliary and the past participle, and different types of subjects (see Table 4.25 above) – could have applied in these corpora as well.

As with my data, auxiliary alternation in pronominals only applied to a limited number of speakers in Sankoff's data: out of a total of 132 speakers, only 10 had any tokens of pronominal verbs conjugated with *avoir* over the three corpora, for a total of 24 tokens (plus one other example from another person present in the interview with Speaker 35 in 1971). Table 4.26 shows a breakdown of the total number of *je m'ai* tokens per speaker.

The symbol 'n/a' indicates that the data is not available for that year[184] and a grey cell that the speaker has not been interviewed that year (see Table 2.3 above for additional details on the longitudinal Montréal French sample).

Speaker ID	N of *je m'ai* tokens by year interviewed			N of *je me suis* tokens by total year(s) interviewed				Total N	% *avoir* selection
	1971	1984	1995	1971	1984	1995	Total		
51	2	1		0	0		0	3	100
63	1			0			0	1	100
37	1*			0			0	1	100
52	3	0		n/a	n/a		1	4	75
33	5			4			4	9	55,6
15	3	1		n/a	n/a		10	14	28,6
96	1			5			5	6	16,7
2	0	3	1	n/a	n/a	n/a	24	28	14,3
75	1	0		n/a	n/a		14	15	6,7
90	n/a	1*	n/a	n/a	n/a	n/a	22	23	4,3

Table 4.26 Speakers with any examples of *je m'ai* (vs *je me suis*) from all three previous Montréal French corpora, by year interviewed, in decreasing order of % *avoir* selection (adapted from Sankoff 2016, personal communication). * = self-correction to *me suis*

Despite the limitations described above, a few qualitative observations can still be made.

One speaker used *avoir* with a pronominal verb in a direct discourse, quoting themselves, as shown in (68).

(68) *Je suis capable de m'excuser à eux-autres, je suis capable de dire que j'ai fait' une erreur, je suis capable de dire « **je m'ai trompé** », je suis capable de dire « tu avais raison ».* (Speaker 2, 1984, par. 317, line 2176)
'I am able to apologize to them, I am able to say that I have made a mistake, I am able to say "I was mistaken", I am able to say "you were right"'

[184] With the data made available to me by Gillian Sankoff, it is not always possible for me to know in which corpora the *je me suis* tokens appear (this caveat applies to speakers 52, 15, 2, 75, and 90).

Another speaker shows auxiliary alternation within the same verb, *se marier* (compare examples 69a-b), but with different subject pronouns: whereas in the same sentence she uses *avoir* in the 1SG with *se marier* but uses *être* in the 3SG with *se remarier*, as illustrated in example 69(a). Example 69(b) also shows auxiliary alternation within the same sentence with different verbs, *se marier* with *être* and *s'acheter* with *avoir*.

(69a) *Puis: l'année ' je* ***m'ai mariée*** *bien maman elle* ***s'est remariée*** *elle.* (Speaker 15, 1971, par. 160, line 331)
'Then the year I got married well mom she got remarried'

(69b) *Quand je* ***me suis mariée*** *là je* ***m'ai t acheté*** *un poêle.* (Speaker 15, 1971, par. 176, line 396)
'When I got married I bought myself an oven'

The epenthetic 'T' present in example 69(b) also appears two other times in examples 70(a-b), all uttered by the same speaker.

(70a) *Ah tabarouette, je* ***m'ai t en-revenue***[185] *avec un mal de tête.* (Speaker 15, 1984, par. 622, line 2107)
'Ah shucks, I came back with a headache'

(70b) *Mais là un beau jour je* ***m'ai t aperçue*** *que: il avait: il avait une maîtresse.* (Speaker 15, 1984, par. 309, line 873)
'But then one fine day I realized that he had, he had a mistress'

Out of the 24 tokens of *avoir* with pronominals, five surfaced with an intervening element between the auxiliary and the past participle.[186] This factor group was found to be highly significant in the multivariate analyses performed on intransitive verb data (see Tables 4.2 and 4.5 above). The elements that appear between the auxiliary and the past participle in the

[185] In Québécois French, the locative adverb *en* always surfaces between the auxiliary verb and the past participle (*Je me suis (t) en allée*) and not between the reflexive clitic and the auxiliary verb, as is the case in standard French (*Je m'en suis allée*).
[186] This information is not available for the equivalent *être* tokens.

previous three corpora of Montréal French are adverbs *déjà* 'already' (illustrated in example 71)*, tellement* 'so'*, bien bien* 'really'*, jamais* 'never', and *pas* 'not'.

(71) *Je: je me dis: je:* ***je m'ai déja tenu*** *avec des: des Français. Des gars qui venaient de France puis: non.* (Speaker 33, 1971, par. 568, lines 1257-1258)
'I, I say to myself, I, I have already hung out with French people. Guys who came from France and no'

Additionally, the corpus of 1971 contains two tokens of false starts with *avoir* followed by self-corrections to *être* uttered by two different speakers. They can be found in (72).

(72a) *Bien: je* ***m'ai: je me suis*** *cassé un bras.* (Speaker 33, 1971, par. 219, line 381)
'So, I, I broke my arm'

(72b) *Là: là j'ai: je* ***m'ai: je me suis lancé*** *dans les affaires.* (Speaker 75, 1971, par. 169, line 289)
'So, so, I, I, I started a business'

In the 2016 corpus, only one probable false start with *avoir* including a reflexive clitic that was followed by a self-correction to *être* was recorded. It is shown in (73).

(73) *Mais je* ***m'ai, suis dit****, j'y ai proposé de peut-être, avec le, pour sa, son secondaire 5, peut-être faire un genre, trouver une couturière, faire faire sa robe de bal de finissant avec ça.* (Linda)
'But I thought to myself, I suggested to her that maybe, with the, for her, her *secondaire 5*,[187] maybe make a sort of, find a seamstress, have her prom dress made out of it'

With the pronominal verbs in 2016, all the other 'false starts' with *avoir* did not include the reflexive clitic (*me, te, s', vous, se,* etc.).

[187] In Québec, it corresponds to the last year of secondary school and is the equivalent of the second-to-last year of high school (grade 11).

It was also possible to search the three earlier corpora for instances of 3PL *plus-que-parfait* subjects (*ils s'avaient*[188]) as well as the masculine 3SG *il s'avait* (but not their *être* counterparts), and one case of each was found, reported respectively in (74a) and (74b). These two tokens have not been included in the distributional analysis presented in Table 4.27 (next page), encompassing only the 1SG *passé composé* tokens.

(74a) *Puis: à part ça bien il savait plus où ce-que: son frère restait hein <non> ça devait faire à-peu-près cinq: ah neuf ans qu'ils* ***s'avaient pas vus.*** (Speaker 72, 1971, par. 335, line 548)
'And on top of that well he did not know where his brother lived anymore <no> it must have been five ah nine years since they had seen each other'

(74b) *Il avait fait ouvert la vitre par son chauffeur puis il* ***s'avait penché*** *puis il nous avait envoyé ça.* (Speaker 32, 1971, par. 198, line 337)
'He had his driver roll down the window and he bent over and he sent it to us'

However, because of unavailable access to the total number of *être* tokens in the entire three corpora (including ones from non-variable speakers), it is not possible to evaluate the overall rate of *avoir* selection with pronominals for the previous three corpora, as was done for the 2016 data in §4.3.1. Nevertheless, with the data available, I calculate the percentage of *avoir* selection for each variable speaker, as shown previously in Table 4.26, and compare it to the rates of the six speakers of the 2016 corpus who used *avoir* with pronominals in the 1SG. Total rates of *avoir* selection with pronominals by speaker from the previous three corpora are outlined again in Table 4.27, in a simplified version of Table 4.26 above.

[188] Laurentian French speakers tend to use the pronoun *ils* 'they.MASC' in the 3PL even for feminine referents (Battye 1991: 272). See examples 38 and 79.

Speaker ID	N of tokens of *je m'ai*	N of tokens of *je me suis*	Total N	% *avoir* selection
51	3	0	3	100
63	1	0	1	100
37	1	0	1	100
52	3	1	4	75
33	5	4	9	55,6
15	4	10	14	28,6
96	1	5	6	16,7
2	4	24	28	14,3
75	1	14	15	6,7
90	1	22	23	4,3
Total	**24**	**80**	**104**	**23,1**

Table 4.27 Percentage of *avoir* selection (*je m'ai* vs *je me suis*) per speaker from all three previous Montréal French corpora in decreasing order of % *avoir* selection (adapted from Sankoff 2016, personal communication)

Among the ten speakers that exhibit auxiliary alternation with pronominal verbs in the previous three corpora, the range of *avoir* rates is considerable (between 4,3% and 100%) and the average rate of *avoir* selection is 23,1%. While the table above suggests that there are three categorical users of *avoir* with 1SG pronominals (Speakers 37, 63, and 51), not enough tokens have been recorded by Sankoff to make such assumptions.

For comparison's sake, the same rates have been calculated for the four speakers who used *avoir* with pronominals in 1SG in *passé composé* in 2016. Their totals can be found in Table 4.28. Speakers Philippe and Julie also each produced two pronominal tokens with *avoir* but not in *passé composé*: Philippe's tokens were both in *plus-que-parfait* and Julie's were both in *conditionnel passé*.

Speaker ID	N of tokens of *je m'ai*	N of tokens of *je me suis*	Total N	% *avoir* selection
Linda	7	1	8	88,9
Jacynthe	1	3	4	40
Denise	1	7	8	9,1
Gaëtan	1	14	15	6,3
Total	**10**	**25**	**35**	**28,6**

Table 4.28 Percentage of *avoir* selection (*je m'ai* vs *je me suis*) by speaker in the 2016 Montréal French corpus

With the data of these four speakers, the total rate of *avoir* selection with pronominals for 2016 with *passé composé* is slightly higher than in 1971 (28,6% vs 23,1%). In terms of distribution, it would appear that approximately the same proportion of speakers used *avoir* with a pronominal verb in *passé composé*: 8,3% of my speakers (4 out of 48) used *avoir* at least once with a pronominal verb, whereas 7,5% of the previous Montréal French speaker samples did so (10 out of 132). Since it is not currently possible to know the total rate of *avoir* selection with pronominals of the previous three Montréal French corpora, it is difficult to make assumptions as to the nature of the variation and/or the direction of change since 1971. Nevertheless, it is plausible that the rate of *avoir* selection with pronominals has decreased alongside *avoir* selection with intransitive verbs.

4.4. Avoidance mechanisms

As stated in §3.2.1.2., two different types of potential avoidance mechanisms were identified, i.e. instances where speakers use various linguistic strategies equivalent in meaning but different in form from periphrastic tenses with auxiliary *être*: the use of *avoir été* in place of *être allé* and the use of historic present rather than the *passé composé*. The tokens of *avoir été* and historic present that have been collected are analysed with *Shiny Rbrul* (Johnson 2017) in order

to determine whether these strategies represent an alternative way of expressing past actions without using the *être* auxiliary.

4.4.1. Uses of *avoir été* in place of *être allé*

In the 2016 corpus, many speakers (26 out of 48) used the lexical verb *être* in contexts where *aller* might be expected, and since the auxiliary that always surfaces with the verb *être* is (invariable) *avoir* in both the standard and the non-standard language, I wanted to test whether this potential avoidance mechanism could be dictated by the same variables as the other intransitive verbs studied here.

In order to determine how the occurrences of *aller* and *être* were distributed, a multivariate analysis was performed on the data of speakers who had shown an alternation between *être allé* and *avoir été*. The three tokens of *aller* that had appeared with *avoir* (see Table 4.9) were removed from this analysis, as well as all the tokens of *ça a bien/mal été* 'it went well/badly' and *ça avait bien/mal été* 'it had gone well/badly' because in my corpus there were no occurrences of *c'est bien allé* ou *c'est mal allé*[189]. The *conditionnel passé* and past infinitives tokens were also excluded from the analysis because there were too few tokens to be tested statistically (four each).

Moreover, the data from speakers who were categorical users of either *être allé* (22 speakers) or *avoir été* (3 speakers) were excluded from the analysis. Interestingly, these last three speakers, Jean-François (20 tokens of *avoir été*), Yves (9 tokens), and Linda (14 tokens), show very different rates of overall *avoir* selection with the other intransitive verbs, 10%, 18,2% and 39,5%, respectively. It is therefore only the data of the remaining 23 speakers that have been included in the statistical analysis, since they show variable use of the two constructions. The results of the multivariate analysis with *Shiny Rbrul* (Johnson 2017) is presented in Table 4.29.

[189] Contrary to what the Académie française suggests in example 30(d) above, these two sentences sound ungrammatical to me, as a native speaker of (Montréal) French.

MODEL 4	**Multivariate analysis of the likelihood of *avoir été* (vs *être allé*)**		
Model formula: Variant ~ Code.name + Intervening.element + Tense + Clause.structure + Type.of.subject			
Input probability	0.146		
Total rate	19,1%		
Total N	446		
R^2	0.481		
Deviance	286.243		
Significant factor groups[190]	**Factor weights**	**% *avoir***	**Total N**
Speaker[191]*******	*p* = 1.11e-18		
Gaëtan	0.977	86,4	22
Philippe	0.889	54,5	11
Alexis	0.874	47,4	19
Marie-Laurence	0.854	50,0	6
Denise	0.839	42,3	26
Marie-Jeanne	0.824	38,1	21
Hugo	0.654	20,0	15
Carl	0.629	20,0	10
Guylaine	0.574	22,2	27
Kim	0.458	10,0	20
Sara	0.396	10,0	10
Sylvain	0.391	7,1	14
Mathieu	0.34	6,7	15
Sophie	0.299	8,0	25
Martine	0.288	6,3	16
Rachel	0.283	5,0	20
Julie	0.283	4,6	22
Jacynthe	0.28	5,6	18
Denis	0.257	13,6	22
Florence	0.218	4,8	21
Annouck	0.161	3,5	29
Caroline	0.126	3,2	31
Johanne	0.115	3,9	26
Tense***	*p* = 3.28e-04		
Plus-que-parfait	0.712	37,0	46
Passé composé	0.288	17,0	400
Type of subject*	*p* = 0.0145		
Relative *qui*	0.923	20	5

[190] The factor groups which did not prove to have a significant effect are 'Polarity of the sentence', 'Gender of the speaker', 'Type of subject', 'Form of auxiliary', 'Person of the subject', 'Age of the speaker (continuous)', 'Age of the speaker (categorical = young vs old)', 'Possibility of transitive use', 'Pronoun before auxiliary', 'Hodiernal action', 'SPS of the speaker'.

[191] The factor group 'Speaker' had to be set as a fixed variable in this model because no mixed-effects model could converge when it was set as a random one. This makes the comparison of Model 4 with Models 2 and 3 less straightforward.

Weak pronoun	0.729	19,4	412
Proper noun	0.39	20	5
Strong pronoun	0.384	10,5	19
Common noun DP	0.0724	20	5
Intervening element (between aux and past participle)*	$p = 0.0152$		
Intervening element	0.661	29,0	31
No intervening element	0.339	18,3	415
Clause structure*	$p = 0.0389$		
Main	0.629	20,2	381
Embedded	0.371	12,3	65

Table 4.29 *Shiny Rbrul* (Johnson 2017) results for Model 4: statistical effects of the significant factor groups on *avoir été* (vs *être allé*), by *p* values and factor weights

Since more than half of the participants are categorical users of one variant or the other, it is clear that the speaker effect on this variable is considerable. But even after removing categorical users from the dataset, Table 4.29 indeed shows that the factor group with the strongest effect determining the use of *avoir été* is indeed the individual speaker, with a very highly significant *p* value of 1.11e-18, regardless of their age, their gender or their SPS. Out of the three categorical *avoir été* users, two belonged to the LOW SPS (Yves and Linda) and the third one (Jean-François) is a HIGH SPS speaker. Near-categorical *avoir été* speaker Gaëtan also belongs to the HIGH SPS. Since Sankoff (2019: 204) reports that the vast majority of the 1971-1984 *être allé* non-users (8 out of their 55 speakers never used *être allé*) belonged to the LOW SPS, it might be the case that we are witnessing a change in progress due to the fact that use of *avoir été* is no longer socially restricted.

The factor group with the second strongest effect is the tense of the verb, with a highly significant *p* value of 3.28e-04. Table 4.29 shows that the occurrences conjugated in *plus-que-parfait* have a much higher chance of surfacing as *avoir été* instead of *être allé* (FW of 0.712) than the ones in *passé composé* (FW of 0.288). This discrepancy between tenses might be explained

by their relative frequency: since the *passé composé* of this verb appears more frequently in speech (almost ten times more often than the *plus-que-parfait* in my corpus), it might therefore be the case that, especially with such a frequent verb as *aller*, it is less prone to modification (Martinet 1969; Bybee 2010).

The factor group with the third strongest effect is the type of subject, with a significant $p = 0.0145$. Relative *qui* and weak pronouns favour *avoir été* with respective FWs of 0.923 and 0.729. Whereas proper nouns, stroung pronouns, as well as common noun DPs all disfavour *avoir été* with FWs of 0.39, 0.384, and 0.0724. In contrast, relative *qui* disfavoured *avoir* use with intransitive verbs in Model 3 (see Table 4.5 above).

The factor group with the fourth strongest effect on *avoir été* (with a significant $p = 0.0152$) is whether there is the presence of an intervening element between the auxiliary verb and the past participle. When intervening material occurs, it favours *avoir* (FW of 0.661) and when there is not it disfavours it (FW of 0.339). This intervening material is usually an adverb, illustrated in (75a), and including negative sentences (where the single negation marker appears between the auxiliary and the past participle), as shown in (75b).

(75a) *Ah,* ***j'ai déjà été*** *à l'école avec Céline* [*Dion*]. (Linda)
'Ah, I once went to school with Céline [Dion]'

(75b) *Euh* ***on a pas été*** *à, voyons, le château, Versailles.* (Guylaine)
'Huh we did not go to, let's see, the castle, Versailles'

The last factor group having a significant effect on *avoir été* that was detected by the multivariate analysis is the clause structure ($p = 0.0389$). When the verb surfaces in a main clause it is more likely to surface as *avoir été* (with a FW of 0.629), and as *être allé* when it is in an embedded one (with a FW of 0.371). While nothing seems to indicate that, so far, we are facing a change in progress, i.e. a projected increased usage of *avoir été* in time, it is the case

that linguistic innovations tend to first take place in main clauses, later spreading to embedded ones, which are more resistant to change (Lightfoot 1991: ch. 3; Harris & Campbell 1995: 27).

4.4.2. Historic present

In this section, I test whether the use of historic present by the speakers can represent another strategy to avoid conjugating verbs with auxiliary *être*. In order to test that hypothesis, I therefore wanted to verify whether the same contexts that favour the use of the *avoir* auxiliary in intransitive verbs are also the ones that favour the surfacing of the historic present for the same verbs. An example of historic present can be found in (76) (see also example tokens in 31(a-b) presented in §3.2.1.2.)

(76) ***J'arrive*** *là, pis ça me tentait pas d'y voir la face à ce gars-là moi.* (Gaëtan)
'I get there, and I did not feel like seeing this guy's face.'

Forty speakers out of 48 produced a total of 308 historic present tokens with 13 *Ê*-verbs (the other eight speakers did not use a single token of historic present).

The multivariate analysis outlining the significant factor groups for the use of historic present instead of *passé composé* is shown in Table 4.30. Only the *passé composé* tokens were included, since the historic present cannot be used in speech as an alternative to the *conditionnel passé* or the *plus-que-parfait*. The eight speakers who never used the historic present were excluded from the analysis, which brought down the total number of tokens to 1920.

MODEL 5	**Multivariate analysis of the likelihood of historic present (vs *passé composé* tokens with *être*)**		
Model formula: Variant ~ Age of the speaker + Clause structure + Hodiernal.recoded + Polarity of sentence + Pronoun.before.aux + (1 | Code.name) + (1 | Verb.collapsed)			
Input probability	0.00445		
Total rate	16%		
Total N	1920		
R^2	0.358		
Deviance	1357.775		
Significant factor groups	**Factor weights**	**% *avoir***	**Total N**
Clause structure* **	*p* = 6.51e-07		
Main clause	0.627	17,7	1556
Embedded clause	0.373	8,8	364
Pronoun before conjugated verb* **	*p* = 2.07e-04		
No pronoun before verb	0.795	16,3	1882
Pronoun before verb	0.205	2,6	38
Hodiernal action **	*p* = 0.00118		
Pre-hodiernal	0.583	17,5	1555
Hodiernal	0.417	9,9	365
Age of the speaker*	*p* = 0.0118		
+1 (year)	Logodds = 0.0298		
Polarity of sentence*	*p* = 0.0217		
Positive	0.655	16,4	1858
Negative	0.345	4,8	62
Speaker (40)	Random		
Lexical item (13)	Random		

Table 4.30 *Shiny Rbrul* (Johnson 2017) results for Model 5: statistical effects of the significant factor groups on the use of historic present, by *p* values and factor weights

Table 4.30 shows that the total rate of historic present use is 16% and that four different linguistic factor groups as well as one social factor group are at play in the variation. The factor group with the strongest effect is clause structure with a very highly significant *p* = 6.51e-07: when the token surfaces in a main clause it favours the use of the historic present, with a FW of 0.627, and when surfacing in an embedded clause it disfavours it, with a FW of 0.373.

The factor group with the second strongest effect (with very highly significant p = 2.07e-04) is whether or not an object or locative pronoun appears before the conjugated verb. The absence of a pronoun before the conjugated verb favours the historic present, with FW of 0.795, and the presence of a pronoun disfavours it, with a FW of 0.205. This is an opposite effect from what was recorded in Models 2 and 3.

The factor group with the third strongest effect on the selection of historic present is whether the action is hodiernal, with a highly significant with a p = 0.00118: when the action of the verb does not take place within 24 hours of the moment of speech, it favours the use of historic present (FW of 0.583) and hodiernal action disfavour it (FW of 0.417). This finding is not entirely surprising given that the very likely use of the historic present in the context of narrations (i.e. the sociolinguistic interviews) serves to make the action described more relevant to the interlocutor and to the moment of speech. One can therefore assume that a speaker describing an event that took place in the previous 24 hours does not feel the need as much to use the historic present than when relating events that are more distant in time.

The factor group with the fourth strongest effect is the age (continuous) of the speakers, with a significant with a p = 0.0118: the older the speaker, the more likely they are to use the historic present (for each additional year the logodd increases by 0.0298).

The last factor group with a significant effect to be identified by *Shiny Rbrul* (Johnson 2017) is the polarity of the sentence with a significant p = 0.0217. A positive sentence favours the use of historic present (FW of 0.655) and a negative one disfavours it (FW of 0.345).

A closer look at the rates of historic present selection of the various verbs,[192] presented in Table 4.31, shows that they vary considerably, ranging from 3,1% to 38,5%.

[192] The lexical effect could not be tested on this data with *Shiny Rbrul* (Johnson 2017) because no model could converge when the lexical item was treated as a fixed effect, rather than a random one.

Verb	N of historic present tokens	Total N of tokens (hist. pres. + *être*)	% historic present
(re)passer	15	39	38,5
(re)descendre	12	33	36,4
(res)sortir	29	96	30,2
arriver	110	393	28
(re)monter	9	33	27,3
(re)rentrer	29	116	25
rester	9	77	11,7
(re)partir	29	252	11,5
(re)tomber	10	91	11
(re)venir	30	309	9,7
déménager	1	14	7,1
retourner	3	49	6,1
aller	22	716	3,1

Table 4.31 Distribution of historic present usage per verb (in decreasing order)

It also appears that different linguistic mechanisms are at play in the choice between historic present and auxiliary *être* when contrasted with the factor groups that constrain *avoir* selection, with the exception of the statistical effect of the factor group 'Hodiernal action'. Though pre-hodiernal actions favour both the use of *avoir* and the use of the historic present in Models 2, 3, and 6, this finding might not be sufficient to say that the use of historic present represents an avoidance strategy. While the data presented above in Tables 4.30 and 4.31 suggest that there may be a lexical and a (non-)hodiernal effect to the variation, it would seem that an avenue worth exploring might be discourse analysis. Since it is highly likely that the use of historic present serves to captivate the attention of the interlocutor, it might be the case that this strategy serves specific pragmatic discourse functions; for example, that certain topics of conversation surface more frequently in the historic present than others, i.e. the Danger of Death question (see §3.1.3.).

4.5. Summary of findings

Overall, the intransitive data yield an average selection rate of *avoir* of 10,7% and the pronominal verbs display an average selection rate of *avoir* of 1,3%. A summary of these results is presented in Table 4.32, which shows the combined rate of *avoir* selection to be 7,4%.

	N of *avoir*	**Total N**	**% *avoir***
Total for intransitive verbs	253	2374	**10,7**
Total for pronominal verbs	17	1263	**1,3**
Total overall	270	3637	**7,4**

Table 4.32 Distribution of *avoir* selection in the Montréal corpus (2016) by verb type

With intransitive verbs, the best-fitting model, Model 3 (Table 4.5 in §4.2.2.3), reveals that the presence of a clitic pronoun before the auxiliary, the inanimacy of subjects, membership to the lowest SPS, an action that is pre-hodiernal, the presence of intervening material between the auxiliary and the past participle, as well as the conditional past tense are all statistical predictors of *avoir* selection. Interestingly, the factor group with the most significant effect (p = 8.31e-06) on my data is whether there is a clitic pronoun before the auxiliary, and this is the first time that such a variable has been tested on auxiliary alternation data. This suggests that any intervening material, regardless of its nature and of its placement with regard to the auxiliary verb, acts as a favouring context for the surfacing of *avoir*.

My qualitative and descriptive analyses also revealed that the effect of individual speakers is important in the auxiliary alternation observed in the Montréal French 2016 data: there was a lot of interspeaker variation and most speakers tended to be categorical users of

one or the other auxiliary with a given verb, at least for the duration of the recording.[193] Distributional analyses have also hinted at the probable effects of the lexical item (§4.2.3.2.) and of varying verb meanings (§4.2.5.2.1.) on auxiliary variation.

An updated study conducted by Sankoff (2019) with the software *R* (*R* Core Team 2004-2019) on the combined 1971, 1984, and 1995 data also confirms most of my results in terms of social and linguistic conditioning. A major difference between her study and mine lies in the influence of age on auxiliary alternation data. While age only seems to be a significant factor group in my pronominal auxiliary alternation data (with older speakers using *avoir* more than younger ones) and not with my intransitive data, the factor group 'Date of birth' did have a statistically significant effect on Sankoff's combined 1971, 1984, and 1995 data ($p < 0.5$), with older speakers using *avoir* more than younger ones (Sankoff 2019: 207).

With pronominal verbs, a very small rate of *avoir* selection of 1,3% was recorded. While no mixed-effects model could be produced because of the small amount of variable data collected, a distributional analysis reveals that 1SG subjects, membership to the lowest SPS, increasing age of the speaker, the presence of a clitic object pronoun before the auxiliary, and the conditional past tense are the best predictors of *avoir* selection with pronominal verbs. A preliminary look at previous data from older corpora of Montréal French (1971, 1984, and 1995), executed with the help of Gillian Sankoff, shows that there is indeed inter- and intraspeaker variation in the auxiliation of pronominal verbs, and that the variation was also restricted to a small number of speakers, and mostly with 1SG subjects. However, insufficient access to the metadata of speakers and to the *être* contexts prevents me from making further real-time observations in terms of social and linguistic conditioning.

Two additional variants have been studied: two avoidance mechanisms (the use of *avoir été* in place of *être allé* and the use of the historic present), since they each represented

[193] This result should still be taken with a grain of salt, as many speakers sometimes produced only one token per verb uttered.

another way of 'saying the same thing'. As to the use of *avoir été* in place of *être allé*, slightly more than half of the speakers of the corpus were categorical users of one variant over the other (*être allé* had 22 categorical speakers and *avoir été* had 3). The multivariate analysis performed in §4.4.1. on the data of the remaining 23 variable speakers reveals that the effect of the individual speaker is the strongest predictor of *avoir* use, followed by the *plus-que-parfait* tense, and when there is an intervening element between the auxiliary and the past participle. With the historic present, the multivariate analysis in Table 4.30 (§4.4.2.) reveals little other than the fact that older speakers were more likely to use it than younger ones, but encourages us to explore the field of discourse analysis to help explain contexts that see speakers preferring historic present over *passé composé*.

The implications of such findings will be discussed in Chapter 6.

Chapter 5. Triangulation of results

I will now proceed to triangulate my findings with two data samples: I first analyse the grammaticality judgements on auxiliary alternation that were collected as part of the sociolinguistics interviews in 2016, and I then examine the data on the acceptability of two sentences with *avoir* that were collected via the crowdsourcing online platform *Français de nos régions* by Avanzi & Thibault in 2017 (Avanzi *et al.* 2016).

5.1. Grammaticality judgement data

Is there a discrepancy between the speakers' willingness to accept *Ê*-verbs (intransitive and pronominal verbs) that surface with auxiliary *avoir* and the speakers' actual use of *avoir* in the interviews? The mismatch between speakers' self-report of using one variant and the actual use of another variant is typical of a covert prestige scenario (Trudgill 1972), where members of a speech community favour the use of non-standard features without being aware that they are orienting to them, and which serves to form a sense of group identity (in informal speech situations).

As described in Chapter 3, the 48 speakers performed a grammaticality judgement test on the acceptability of auxiliary *avoir* in sentences with the *Ê*-verbs studied, as well as with various pronominal verbs (see Appendix 3C for the full list of sentences). They were asked to listen to 60 short sentences being read aloud to them by me and to say whether they could have potentially heard these sentences being uttered 'naturally' by native speakers of Montréal French. As stated in §3.6.1., it was decided that it would be preferable not to study self-reporting judgements with my participants because morphosyntactic variables are often quite socially marked and phonologically salient, and there was a risk that most of them would deny

using the *avoir* variant. While my speakers were not asked directly whether they would themselves say the sentences of the task, it is still possible to test whether *avoir* selection with certain verbs is a linguistic phenomenon of which they are aware.

In the analysis presented here I first seek to establish whether the hierarchy of verbs based on their acceptability with *avoir* correlates with the results obtained from the sociolinguistic interview data. Second, I verify whether there is a mismatch between the grammaticality judgements of the speakers and their actual usage, and ascertain whether the speakers who rarely use *avoir* still recognize that it is a feature of the variety of French of which they are native speakers. Thirdly, I seek to correlate the various sociolinguistic factors that were significant predictors of *avoir* selection in the corpus with the acceptability of *avoir* in the grammaticality judgement tests.

5.1.1. Grammaticality judgements compared with production data by verb

For each verb under study, I calculated how many speakers out of 48 judged it to be grammatical when conjugated with *avoir*. This yielded an acceptability rate of the *avoir* variant for each verb.[194] Four different groupings of verbs were tested: the intransitive verbs studied in §4.2., but also the 'categorical' *Ê*-verbs (*mourir, naître, devenir*), the rarer *Ê*-verbs (*survenir, parvenir, intervenir*) mentioned in §2.1.2., and pronominal verbs.

[194] In this case, it was not possible to use the magnitude estimation method developed by Bard, Robertson, & Sorace (1996), where the speakers indicate whether a sentence is more or less acceptable than another one. Because this method would have created 48 different relative (rather than absolute) acceptability scales with varying levels of acceptability according to each speaker and would have necessitated additional technical tools to randomize the order in which the prompts were presented to the participants and to analyse the resulting data, because the analysis of grammaticality judgements was partially beyond the main scope of this project, and because of time constraints during the interview, it was judged that it was better to use a more straightforward methodology for this project.

5.1.1.1. Intransitive Ê-verbs

Most intransitive verbs were judged by the vast majority of speakers to be grammatical when used with *avoir*. Only one 'control' sentence, with *venir*, with the verb conjugated with *être* failed to yield a 100% rate of acceptability.[195] The overall results of the acceptability of the *avoir* intransitive variants are presented in Table 5.1. They are then compared to the actual selection rate with *avoir* of these verbs in the production data from the sociolinguistic interviews.

[195] The sentence with the verb *venir,* which was not found to be acceptable by five speakers, is the following:

> *Est-ce que c'est vous qui êtes venu(e.s) ce matin ?*
> 'Are you (plural or formal singular) the one who came this morning?'

It is very plausible that for these five speakers the *est-ce que* form sounded too formal and that they would have preferred a sentence with the interrogative particle *-tu*, as in *C'est-**tu** vous qui êtes venu(e.s) ce matin?*.

Lexical item	Sentences	Total N of 'Yes' out of 48 answers	% of acceptability with *avoir*	% *avoir* selection in interviews
passer	*Les vacances sont déjà finies… Ça* ***a*** *tellement* ***passé*** *vite !*[196]	48	**100**	**58,6**
déménager	*On* ***a déménagé*** *à Sherbrooke y'a deux ans.*	47	**98**	**80,3**
rentrer dedans	*(– C'est de ta faute l'accident d'auto ?)* *– Ben non, c'est lui qui m'****a rentré dedans*** *!*	47	**98**	**89,4**
sortir avec	*J'****ai sorti avec*** *elle pendant 2 ans à peu près.*	47	**98**	**53,3**
demeurer	*Ils* ***ont demeuré*** *à Outremont pendant 15 ans.*	46	**96**	**n/a**
descendre	*Elle a des gros problèmes de genou, je crois pas qu'elle* ***aurait descendu*** *au sous-sol sans sa canne.*	45	**94**	**29,7**
sortir	*C'était trop dangereux, donc j'****ai sorti*** *de là le plus vite que j'ai pu.*	45	**94**	**13,3**
monter	*J'****ai monté*** *jusqu'en haut de la montagne.*	44	**92**	**35,1**
rester	*On* ***a resté*** *neuf jours à Paris, puis après on a pris le train pour Londres.*	44	**92**	**19,4**
retomber	*J'ai comme eu l'impression d'****avoir retombé*** *en enfance !*	44	**92**	**0**
remonter	*Une chance que la bourse* ***a remonté*** *!*	43	**90**	**55,6**
retourner	*On* ***a retourné*** *en arrière parce qu'on avait oublié le chien à la maison.*	43	**90**	**12,1**
tomber	*Durant la tempête de verglas, l'arbre* ***a tombé*** *sur les fils électriques.*	42	**88**	**33,3**
rentrer	*J'****ai rentré*** *par la fenêtre parce que j'avais oublié mes clés.*	41	**85**	**8,9**
partir	*J'****ai parti*** *de là dès que je l'ai vu arriver.*	37	**77**	**2,1**
arriver	*On* ***a arrivé*** *presque en même temps qu'eux.*	16	**33**	**0,6**
revenir	*Elle* ***a revenu*** *chez ses parents après ses études.*	15	**31**	**1,3**
venir	*J'****ai venu*** *aussitôt que j'ai appris la mauvaise nouvelle.*	12	**25**	**0**
aller	*J'****ai allé*** *au cinéma hier.*	8	**17**	**0,4**

Table 5.1 Ranking of the intransitive *Ê*-verbs in decreasing order of acceptability with *avoir* according to the 48 speakers

[196] See Appendix 3C for the full list of sentences alongside their English translations.

The verb *passer* which surfaced with *avoir* in 58,6% of the tokens during the interviews is the only verb that was judged to be grammatical with *avoir* by all the participants. It might be the case that the sentence was readily accepted by the participants because of its inanimate subject (*ça*) and because of the intervening element (the adverb *tellement*), two of the predictors for *avoir* use identified by *Shiny Rbrul* (Johnson 2017) (see §4.2.2.2.).

Table 5.1 also confirms the finding presented in §4.2.5.2.1., concerning lexicalized *rentrer dedans* and *sortir avec*. *Rentrer dedans*, which appeared with *avoir* in 89,4% of cases in the interview data (compared with only 8,9% of *rentrer* tokens), was judged to be grammatical with *avoir* by slightly more speakers than *rentrer*: 47 vs 41. A similar, but much weaker, pattern can be observed for the verb *sortir avec* compared with *sortir*. *Sortir avec* appeared with *avoir* in 53,3% of cases during the interviews (vs 13,3 % for *sortir*), and was judged to be grammatical with *avoir* by just slightly few more speakers than for its core meaning: 47 vs 45 speakers.

The verbs *descendre, monter, rester, retomber, remonter, retourner,* and *tomber* yielded high rates of acceptability with *avoir*, all between 85% and 94%. Though it is possible that the different verb tenses used for the verbs *descendre* and *retomber*, the past conditional and the past infinitive respectively, might have conditioned the speakers to accept these verbs with *avoir* more easily, since these two tenses tend to favour *avoir* use, as explained in §4.2.2.2. and §4.2.2.3.

The verb *partir* particularly stands out because of the gaps that separate it from the verbs that are immediately more/less acceptable with *avoir*. *Partir* surfaced with *avoir* in 2,1% of the tokens during the interviews and was considered to be less grammatical with *avoir* than *rentrer,* with a 77% acceptability rate. However, *partir* was found to be much more acceptable with *avoir* than the next verb, *arriver*, which yielded an acceptability rate of 33%.

Table 5.1 also shows that *arriver, (re)venir,* and *aller* are by far the least acceptable verbs with *avoir* in these sentences, with rates ranging from 17% to 33%. However, while *venir* did not surface once with *avoir* in a total of 174 tokens, it is still considered to be acceptable with *avoir* by 25% of speakers.

Figure 5.1 illustrates the contrast between the rates of acceptability with *avoir* per verb and their actual selection rates with *avoir* in the sociolinguistic interviews, ranked in decreasing order of acceptability.

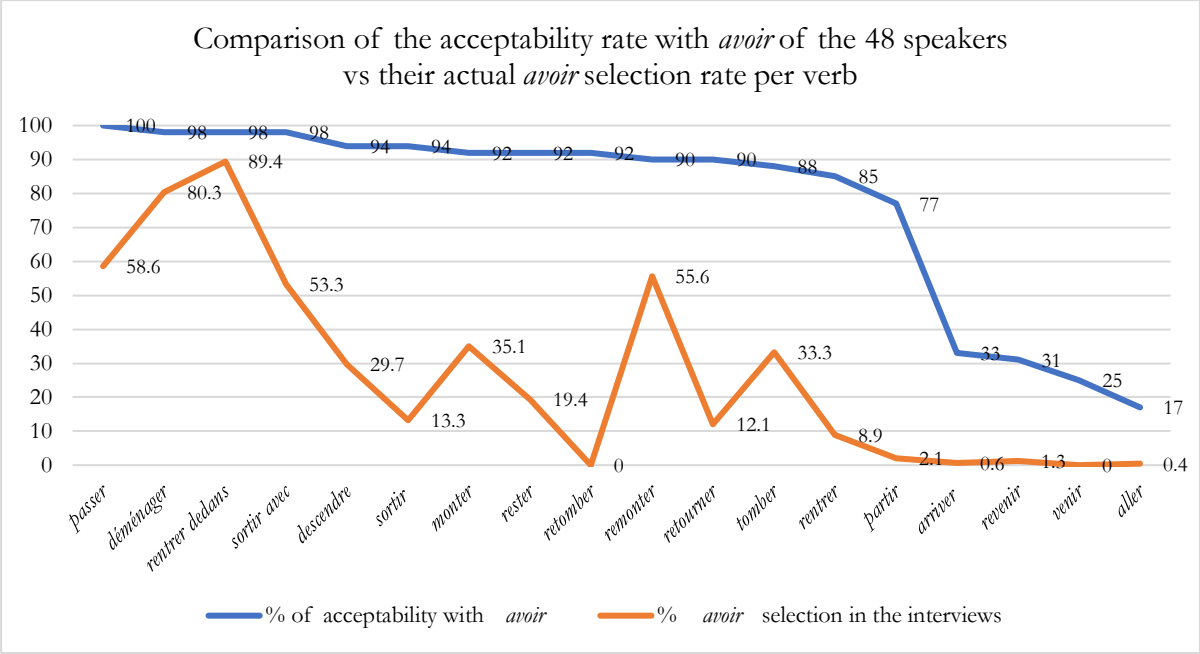

Figure 5.1 Acceptability rates with *avoir* of the 48 speakers compared to their actual *avoir*-selection rates (in percentage, per verb)

Figure 5.1 shows that the two lines do roughly follow a descending slope, especially in the extremes where the patterns match up, but there is nevertheless a lot of variability in the middle portion of the graph. With the verb *remonter*, which slightly stands out from the rest, it seems that a few participants judged the sentence with *avoir* to be unnatural because of its content rather than because of its structure, since it transpired that they were not familiar with the concept of the Stock Exchange and its ebbs and flows.

In order to verify whether this potential correlation could be statistically measured, the above data have been re-plotted against each other. The resulting scatterplot is presented in Figure 5.2, where a regression line (with an R^2 score) has also been fitted.

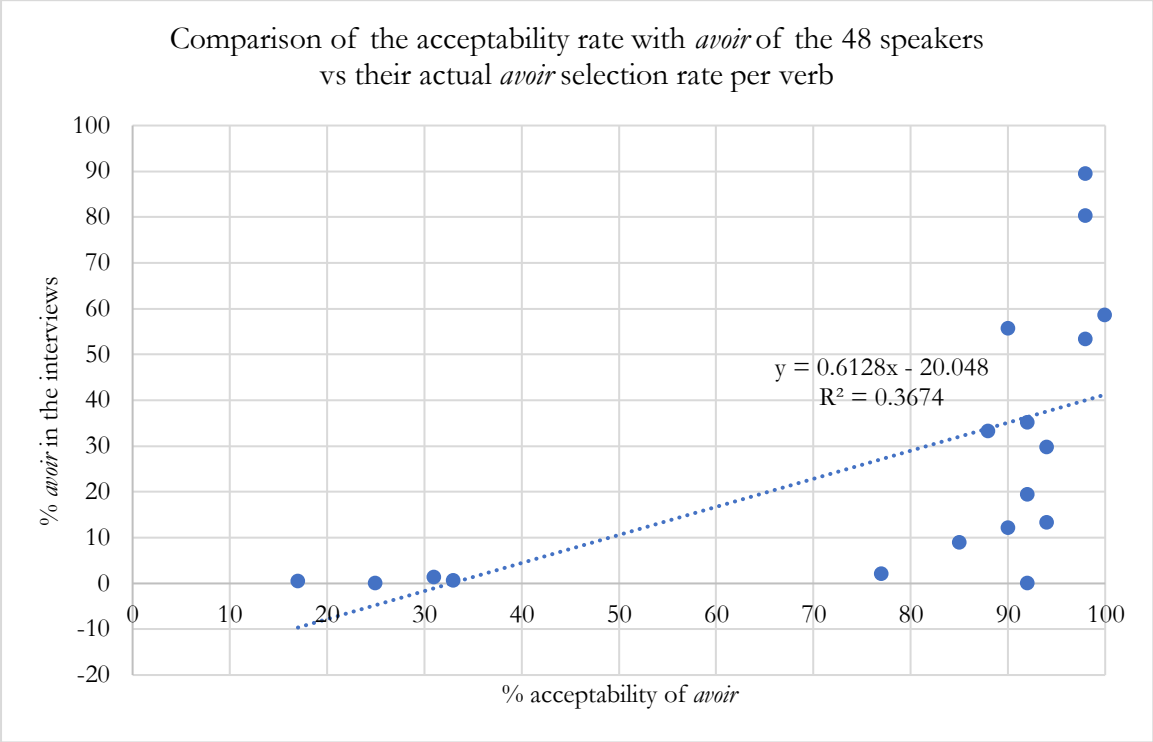

Figure 5.2 Scatterplot of the acceptability rates with *avoir* of each verb compared to their actual *avoir*-selection rates in the interviews (in percentage)

From Figure 5.2, it is possible to notice that it is difficult to fit a regression line ($R^2 = 0{,}3674$) because of the variability around the line, particularly where acceptability is high but production is very variable (to the right of the graph). The first grouping comprises verbs which are very rarely conjugated with *avoir* and are less likely to be accepted with *avoir*: *aller, venir, revenir,* and *arriver*, which had low acceptability rates, between 17% and 33%. The second grouping comprises verbs which yielded very different *avoir* rates during the interviews

(between 8,9% and 89,4%) but were all readily accepted with *avoir* in the grammaticality judgement tests (between 77% and 100% of the speakers).

Overall, it is very possible that some of these sentences with *avoir* were made 'more grammatical' to the speakers, i.e. were more readily accepted with *avoir,* because they contained some linguistic predictors (other than the lexical item) determined in §4.2.2., such as inanimacy of the subject, clitic pronoun before auxiliary, intervening element between the auxiliary and the past participle, past conditional tense, and positive polarity of the sentence.

5.1.1.2. 'Categorical' intransitive Ê-verbs

The verbs *devenir, mourir,* and *naître*, showed no auxiliation variation during the interviews and always selected *être.* As expected, with categorical *Ê*-verbs all the control sentences with *être* yielded a 100% rate of acceptability. The results for the acceptability of *avoir* with categorical *Ê*-verbs are presented in Table 5.2.

Lexical item	**Sentences**	**Total N of 'Yes' out of 48 answers**	**% of acceptability with *avoir***	**% *avoir* selection in interviews**
devenir	*Il* ***a devenu*** *tellement fâché quand il a appris ce qui s'est passé.*	18	**38**	**0** (0/42)
mourir	*Mon grand-père* ***a mouru*** *le mois passé.*[197]	9	**19**	**n/a**
naître	*J'****ai né*** *à Montréal en 1982.*	3	**6**	**n/a**

Table 5.2 Ranking of 'categorical' *Ê*-verbs (*devenir, mourir*, *naître*) in decreasing order of acceptability with *avoir* according to the 48 speakers

[197] Despite potentially introducing a possible confound, it was decided to include this non-standard form of the past participle of *mourir* in the task 1) because, as a native speaker, it was my intuition that this form would have been more acceptable than *j'ai mort,* 2) because I wanted to verify whether that form was still recognized as grammatical by speakers, and 3) because I wanted to avoid a situation where the speakers could have mistaken the past participle *mort* 'dead' for the substantive *la mort* 'death'.

The verb *devenir* was judged to be acceptable with *avoir* by 18 speakers, even though it was not conjugated once with *avoir* during the sociolinguistic interviews (out of 42 tokens). With the verb *mourir,* the use of auxiliary *avoir* combined with the past participle modelled on the simple past (*mouru* vs *mort*) is attested in the language of lower class speakers,[198] as reported by Frei (1929: 86) in his *Grammaire des fautes*. Frei wrote that this form can be found but "*seulement dans le plus bas peuple, complètement inculte*" ('only by people of the lowest class, completely uneducated'). This form was judged acceptable by 19% of speakers. While no tokens of *avoir mouru* were collected during the interviews, one token of *être mouru* was recorded:

(77) *Ah, ils* [*les fleurs*] ***sont mouru(e)s***. (Julie)
'Ah, they [the flowers] have died'

With the verb *naître,* the form with *avoir* was only judged to be acceptable by three speakers. However, it is likely that a modified form of the past participle modelled on the simple past tense of the verb (as is the case in *mourir*), such as *naquis,* would have yielded a higher rate of acceptability.[199]

5.1.1.3. Rare intransitive Ê-verbs

The verbs in this category, namely *intervenir, survenir,* and *parvenir*, were not studied in the interview data because they never surfaced. Moreover, the only 'control' *être*-sentence in the task for these verbs was with the verb *survenir* and it only yielded a 35% rate of acceptability. This sentence is found in (78).

(78) *Il* ***est survenu*** *à l'improviste.*
'He appeared without warning'

[198] It is not because the *passé simple* is still being used by these speakers, but because this past participle is long-standing.
[199] I asked a few native speakers in my circle whether they agreed with my intuition, and they all confirmed that *j'ai naquis* sounded more grammatical than *j'ai né.*

The fact that only 17 speakers out of 48 judged the sentence above to be felicitous indicates that it is likely that the combination of the rare verb *survenir* with the expression *à l'improviste* sounded too formal in this task to be accepted by speakers.[200] The overall results of the acceptability of *avoir* with the three rare *Ê*-verbs are presented in Table 5.3.

Lexical items	**Sentences**	**Total N of 'Yes' out of 48 answers**	**% of acceptability with *avoir***
intervenir	*Elles* ***ont intervenu*** *deux fois pendant le débat.*	42	**88**
survenir	*L'accident* ***a survenu*** *sans qu'on puisse faire quoi que ce soit.*	31	**65**
parvenir	*Il* ***a parvenu*** *à le convaincre d'arrêter de fumer.*	28	**58**

Table 5.3 Ranking of rarer intransitive *Ê*-verbs (*intervenir, survenir, parvenir*) in decreasing order of acceptability with *avoir* according to the 48 speakers

Table 5.3 shows that the verb *intervenir* is more readily accepted with *avoir* (88%) than the other two verbs. It might be the case that the *survenir* sentence with *avoir* (65%) was more easily accepted than its 'control' sentence with *être* (35%) because its subject was inanimate, *l'accident* (see Table 4.2 above). The sentence with *parvenir* was judged to be the least grammatical by the speakers, with an acceptability rate of 58%.

[200] After hearing that prompt, some speakers repeated "'*À l'improviste*?' Non.", seemingly rejecting the expression as possibly being uttered by someone speaking informally on the street.

5.1.1.4. Pronominal verbs

In order to keep the grammaticality judgement task fairly short, no control sentences with *être* for pronominal verbs were included.[201] The overall results of the acceptability of *avoir* with pronominal verbs are presented in Table 5.4.

Lexical item	Sentences	Total N of 'Yes' out of 48 answers	% of acceptability with *avoir*	% *avoir* selection in interviews
se brosser	*Tu peux pas m'embrasser sans* ***t'avoir brossé***[202] *les dents !*	45	**94**	**0** (0/4)
se fouler	*J'avais vraiment mal parce que je* ***m'avais foulé*** *la cheville en tombant.*	40	**83**	**50** (1/2)
se déguiser	*Je* ***m'avais déguisé*** *en Père Noël pour les enfants.*	35	**73**	**100** (1/1)
se mettre	*Je* ***m'ai mis à l'aise*** *quand j'ai vu que personne s'était habillé chic.*	30	**63**	**0** (0/10)
s'habiller	*Après avoir pris ma douche, je* ***m'ai habillé*** *en vitesse.*	30	**63**	**0** (0/5)

Table 5.4 Ranking of the pronominal verbs in decreasing order of acceptability with *avoir* according to the 48 speakers

Table 5.4 shows that even though pronominals were very rarely conjugated with *avoir* during the interviews (an *avoir*-selection rate of 1,3%, see §4.3.1.), their acceptability rates with *avoir* are relatively high, ranging from 63% to 94%. It is noteworthy that the highest acceptability rate with *avoir* was yielded by the sentence that contained the past infinitive (94%), followed by two sentences in the *plus-que-parfait* (73%-83%) and two in *passé composé* (63%). This hierarchy of acceptability of *avoir* with pronominal verbs by verb tense is consistent with the statistical results presented in the distributional analyses (see §4.3.2.), where the ranking of

[201] It was assumed that prompts containing pronominal verbs conjugated with auxiliary *être* were extremely likely to yield a 100% acceptability rate since they surface so rarely with auxiliary *avoir* in natural speech.
[202] The prompt with a past infinitive introduces a confound but since it was taken straight from the Rea 2014 interviews it was decided to include it in the task without any modification, as mentioned in §3.6.1. The sentence *Tu* ***t'as*** *pas* ***brossé*** *les dents* in the *passé composé* may have yielded different responses.

constraints for the factor group 'Tense' sees the *plus-que-parfait* as disfavouring *avoir* less (rate of *avoir* of 42,9%) than the *passé composé* (18,9%).

5.1.2. Grammaticality judgements compared with production data by speaker

For each individual speaker, I calculated how many sentences out of the 18 intransitive sentences with *avoir* and the five pronominal sentences with *avoir* were judged to be grammatical. This yielded an acceptability rate of the *avoir* variants for each speaker.

5.1.2.1. Intransitive verbs

The intransitive verbs under study in this section are the same as those analysed in §5.1.1.1. except for *demeurer*, which only surfaced twice during the interviews. The results of the acceptability of *avoir* per speaker, with the 18 intransitive sentences, are presented in Table 5.5, together with the speakers' actual selection rate with *avoir* in the sociolinguistic interviews.

Speaker pseudonym	Total 'Yes' out of 18 sentences with *avoir*	% of acceptability with *avoir*	% *avoir* selection in interviews
Linda	18	100	39,5
Sylvain	18	100	15,7
Philippe	17	94	45,8
Steve	17	94	10,4
Jean-François	17	94	10
Marie-Laurence	17	94	4,8
Denis	17	94	2,6
Madeleine	16	89	8,5
Martine	16	89	5,6
Hugo	16	89	2,8
Maxime	15	83	15,6
Sara	15	83	14,7
Joël	15	83	14,6
Jacynthe	15	83	12,5
Alexis	15	83	9,3
Julie	15	83	9,3
Kim	15	83	6
Caroline	15	83	5,4
Florence	15	83	2,9
Paul	15	83	1,8
Francis	15	83	0
Yves	14	78	18,2
Dominic	14	78	10,6
Marc	14	78	7,8
Annie	14	78	6,2
Jessica	14	78	5,8
Justin	14	78	5,3
Denise	14	78	5,2
Gaëtan	13	72	13,4
Nathan	13	72	8,3
Martin	13	72	5,7
Johanne	13	72	5,6
Sabrina	13	72	5,5
Guylaine	13	72	5,1
Mathieu	13	72	4,4
David	13	72	4,3

Richard	13	72	2,4
Carl	13	72	1,7
Mario	13	72	1
Sophie	12	67	13,7
Annouck	12	67	4,7
Amélie	12	67	1,7
Christine	11	61	1,2
Rachel	10	56	11,1
Charles-Antoine	10	56	4,4
Marie-Lou	10	56	0
Virginie	8	44	7,9
Marie-Jeanne	8	44	5,7

Table 5.5 Ranking of the 48 speakers in decreasing order of their willingness to accept the 18 intransitive sentences with *avoir*

Table 5.5 shows that there is a great deal of variability in how willing the speakers are to recognize that *avoir* selection is part of their variety of French, even though the vast majority of *avoir* sentences of the task had been uttered by native speakers of Montréal French just three years prior (Rea 2014). The acceptability rates range from 44% to 100% of the verbs/sentences. There is also a lot of variability in the acceptance rates as well as in how they compare with production rates. Figure 5.3 illustrates the contrast between the rates of acceptability of the 18 intransitive verbs with *avoir* of each speaker and their *avoir*-selection rates produced during the sociolinguistic interviews.

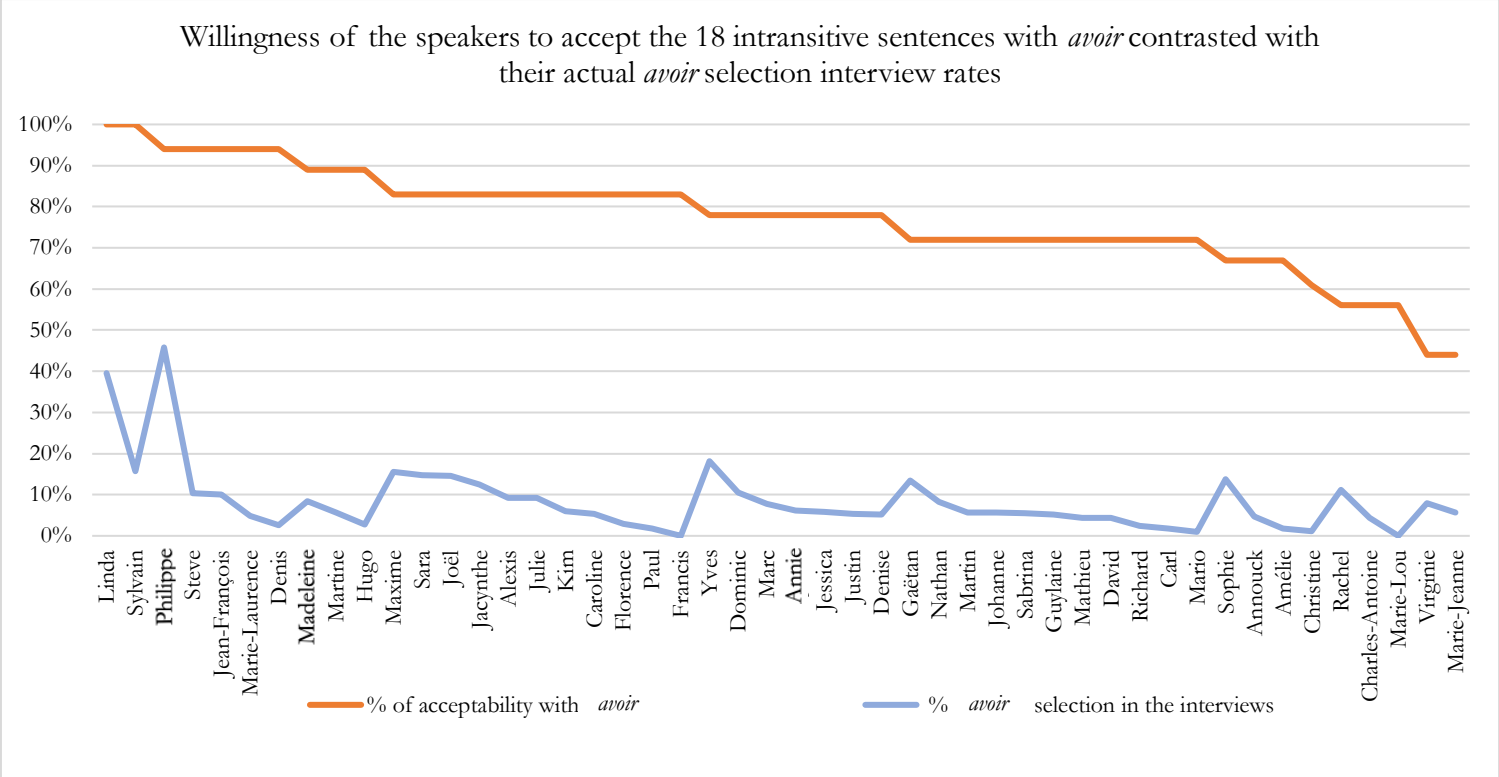

Figure 5.3 Ranking of the 48 speakers in decreasing order of their willingness to accept the 18 intransitive sentences with *avoir* contrasted with their actual *avoir* selection interview rates

In Figure 5.3, the orange and blue lines do roughly follow the same, almost parallel, descending slope, suggesting at first glance that acceptability of *avoir* correlates with actual use of *avoir*, apart from a few peaks caused by speakers Philippe, Maxime, Yves, Gaëtan, Sophie, Rachel, and Virginie. Figure 5.3 also shows that people who do not really use the *avoir* variants in their own speech are aware that they can be found in the variety of French that they speak, even though they seem to find them a bit less grammatical than the speakers who use them more frequently.

In order to verify whether the two lines do exhibit similar slopes, i.e. whether there is a statistical correlation, the data from Figure 5.3 were re-plotted against each other. The resulting scatterplot is presented in Figure 5.4, where a regression line (with an R^2 score) has also been fitted.

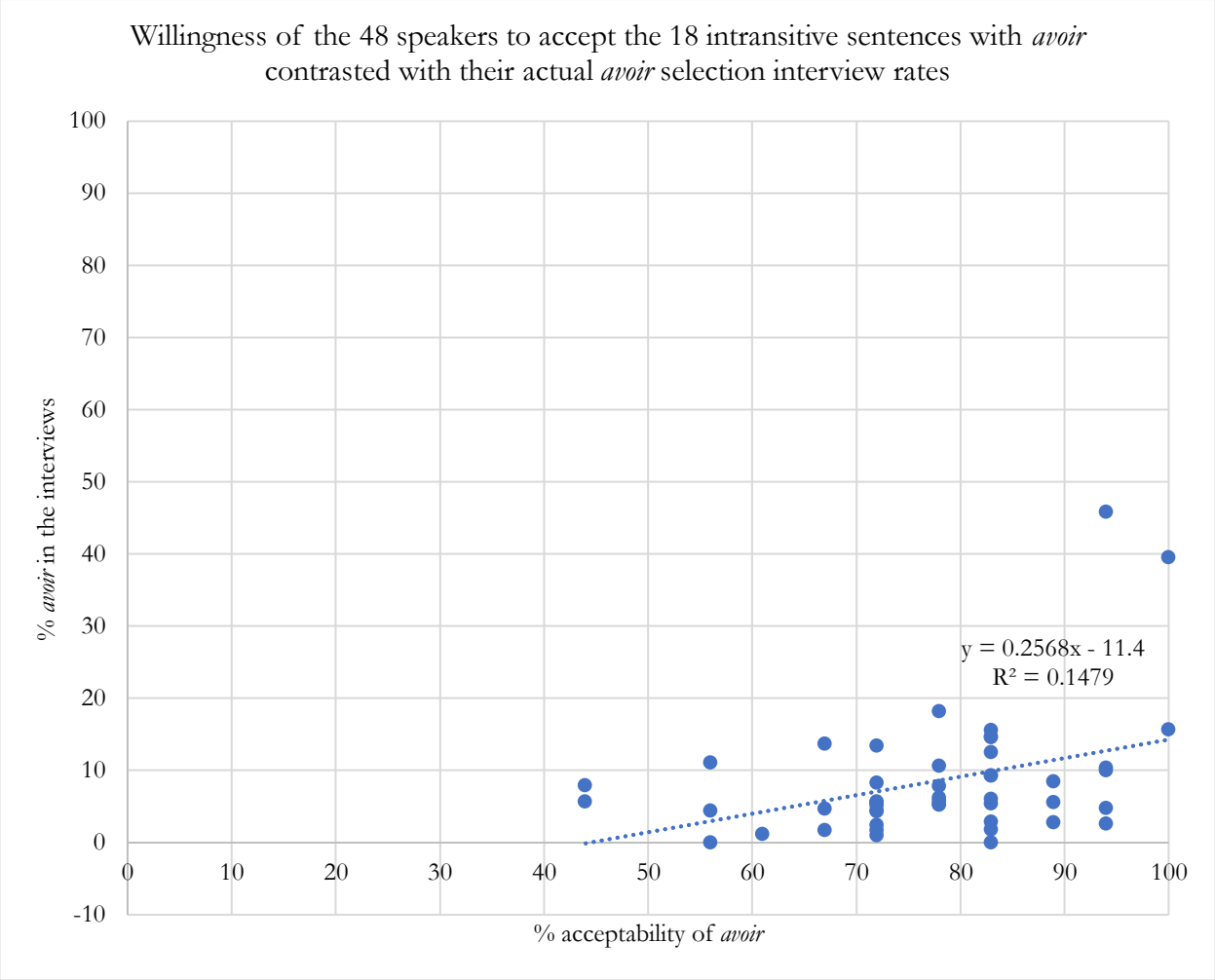

Figure 5.4 Scatterplot of the 48 speakers' willingness to accept the 18 intransitive sentences with *avoir* contrasted with their actual *avoir* selection interview rates

From Figure 5.4, it is clear that it is difficult to fit a regression line (R^2 = 0,1479) because there was a lot of variability in terms of the actual *avoir* rates of the speakers during the interviews (between 0% and 45,8%), as well as in their *avoir* rates resulting from their grammaticality judgement tests (between 44% and 100%). Two LOW SPS speakers, Philippe and Linda, by far the highest users of *avoir* during their interviews (rates of *avoir* selection of 45,8% and 39,5%, respectively), are outliers and removing them from the dataset yields a new linear regression in Figure 5.5.

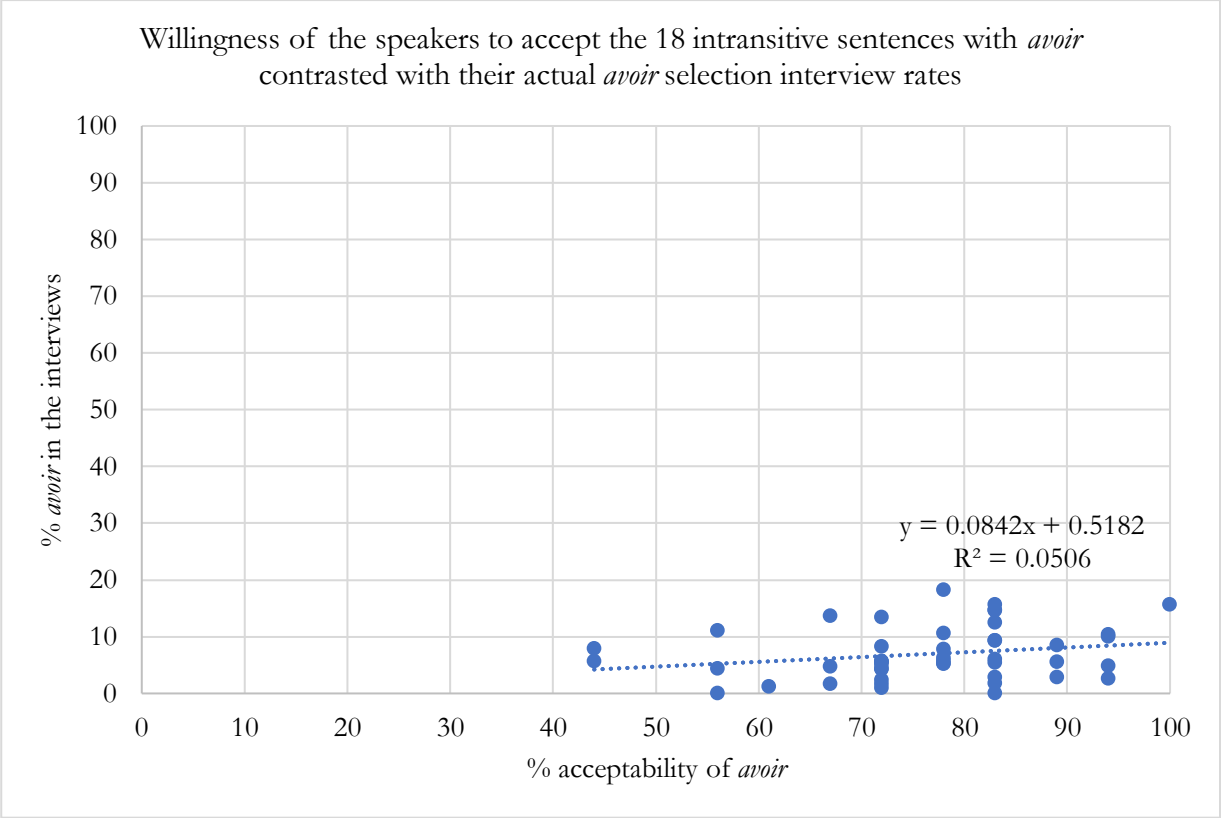

Figure 5.5 Scatterplot of the 46 speakers' willingness to accept the 18 intransitive sentences with *avoir* contrasted with their actual *avoir* selection interview rates

The resulting flat regression line in Figure 5.5 partially confirms my results: some speakers who are not frequent users of the *avoir* variants in their own speech are aware that those can be found in the variety of French that they speak (even though they seem to find them slightly less grammatical than speakers who use them more frequently, like speakers Philippe and Linda for example). Moreover, the very low R^2 score (R^2 = 0,0506) indicates that there is a great deal of variability in how the speakers use *avoir* and that their use of *avoir* is not at all a predictor of their acceptability of it.

5.1.2.2. Pronominal verbs

The results for the acceptability of *avoir* per speaker, with the five pronominal sentences, are presented in Table 5.6, compared to the speakers' actual selection rate of pronominals conjugated with *avoir* in the sociolinguistic interviews.

Speaker pseudonym	**Total 'Yes' out of 5 sentences with *avoir***	**% of acceptability with *avoir***	**% *avoir* selection in interviews**
Linda	5	100	38,1
Philippe	5	100	28,6
Jacynthe	5	100	25
Denise	5	100	3,7
Carl	5	100	0
Christine	5	100	0
Denis	5	100	0
Dominic	5	100	0
Florence	5	100	0
Francis	5	100	0
Hugo	5	100	0
Annie	5	100	0
Jean-François	5	100	0
Joël	5	100	0
Johanne	5	100	0
Kim	5	100	0
Madeleine	5	100	0
Marie-Laurence	5	100	0
Martine	5	100	0
Nathan	5	100	0
Richard	5	100	0
Sara	5	100	0
Steve	5	100	0
Sylvain	5	100	0
Yves	5	100	0
Alexis	4	80	0
Annouck	4	80	0
Caroline	4	80	0
Marc	4	80	0

Maxime	4	80	0
Paul	4	80	0
Guylaine	3	60	0
Jessica	3	60	0
Justin	3	60	0
Martin	3	60	0
Mathieu	3	60	0
Gaëtan	2	40	2,5
Charles-Antoine	2	40	0
Mario	2	40	0
Sophie	2	40	0
Julie	1	20	4,3
Amélie	1	20	0
David	1	20	0
Marie-Jeanne	1	20	0
Marie-Lou	1	20	0
Rachel	1	20	0
Sabrina	1	20	0
Virginie	1	20	0

Table 5.6 Ranking of the 48 speakers in decreasing order of their willingness to accept the five pronominal sentences with *avoir*

Table 5.6 shows that while only six speakers (Linda, Philippe, Jacynthe, Denise, Gaëtan, and Julie) conjugated a pronominal with *avoir* during the sociolinguistic interviews, almost half of the sample, i.e. 21 speakers, accepted all the pronominal sentences with *avoir* even though they had not themselves conjugated a pronominal verb with *avoir* during their interview. For these speakers, the willingness to recognize the grammaticality of *avoir* with a pronominal therefore does not translate at all into an increased usage of that auxiliary verb. This is clear in Figure 5.6, which shows the contrast between the rates of acceptability of the five pronominal sentences with *avoir* of each speaker and their actual *avoir*-selection rates with all pronominal tokens produced in the sociolinguistic interviews.

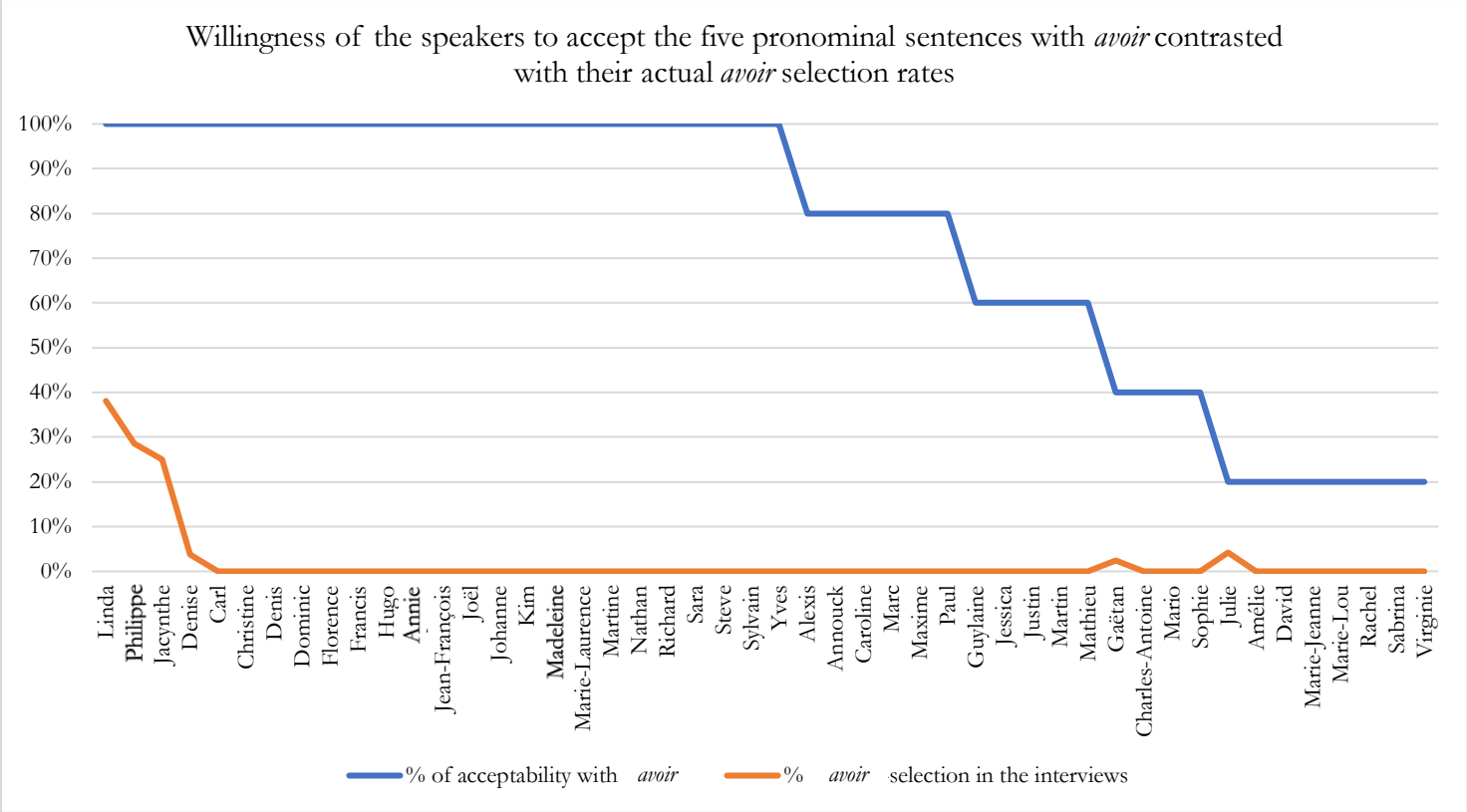

Figure 5.6 Ranking of the 48 speakers in decreasing order of their willingness to accept the five pronominal sentences with *avoir* contrasted to their actual *avoir*-selection rates

Figure 5.6 confirms that the all 48 speakers do accept pronominal verbs conjugated with *avoir* (at least one pronominal sentence out of five) as being a feature of the variety of French spoken in Montréal. As opposed to the strong variability observed in the perceived grammaticality of intransitive verbs with *avoir,* this seems to indicate that most speakers are aware that pronominal verbs conjugated with *avoir* exist in Montréal French and that this judgement is not at all constrained by their respective usage of this variant. Moreover, of the eight speakers who only accepted one sentence, all but one (Marie-Lou) picked the one which contained the past infinitive token, reproduced in (79).

(79) *Tu peux pas m'embrasser sans* ***t'avoir brossé*** *les dents !*
'You can't kiss me without having brushed your teeth'

In Figure 5.6 (and from Table 5.6 above), it is also possible to see that speakers Gaëtan and Julie (the two small peaks protruding to the right-hand side of Figure 5.6) are the only speakers who have used *avoir* with a pronominal verb during their sociolinguistic interview but who did not judge all pronominal sentences with *avoir* in the task to be grammatical.

For consistency, the data from Figure 5.6 were re-plotted against each other. The resulting scatterplot is presented in Figure 5.7, and a regression line fitted.

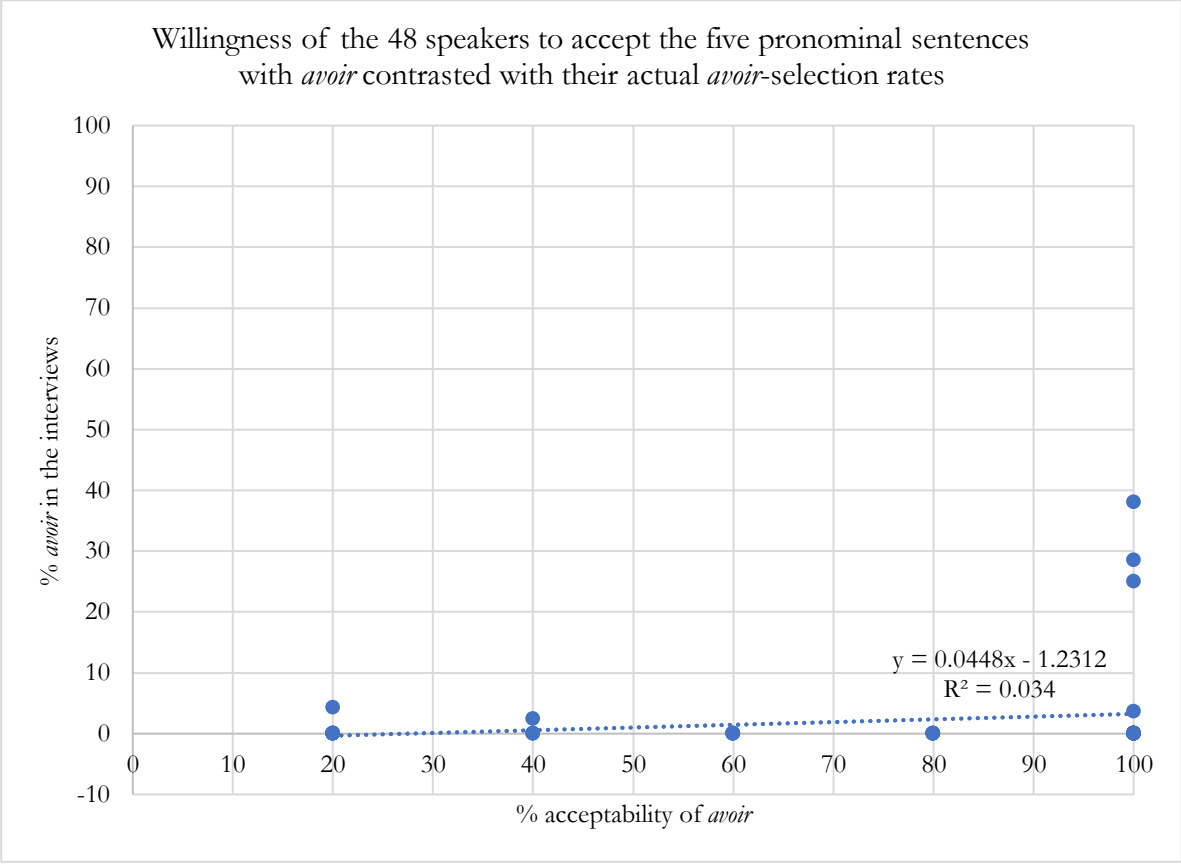

Figure 5.7 Scatterplot of the 48 speakers' willingness to accept the five pronominal sentences with *avoir* contrasted with their actual *avoir*-selection rates

The near-flat regression line confirms my results: all speakers accept pronominal verbs conjugated with *avoir,* irrespective of their usage of this variant. Moreover, the very low R^2

score (R^2 = 0,034) indicates that there is a great deal of variability in how the speakers use *avoir* with pronominals and that their use of *avoir* is not at all a predictor of their acceptability of it.

5.1.3. Correlation with sociolinguistic variables

Since it has been shown that the vast majority of speakers accept most *avoir* variants, both intransitive and pronominal, as part of Montréal French regardless of whether they use them in their own speech, in this section I verify whether the various sociolinguistic factor groups that were identified as predictors by *Shiny Rbrul* (Johnson 2017) in Table 4.2 (see §4.2.2.2.) and Table 4.5 (see §4.2.2.3.), as well as in Table 4.25 (see §4.3.2.) can be held responsible for the disparity in *avoir* acceptability rates recorded in intransitive and pronominal sentences in the grammaticality judgement tests. As stated earlier, in §5.1.1.1. and §5.1.1.2., the difference in acceptability rates for both intransitive and pronominal verbs could be due to a lexical effect and/or other linguistic variables for which I could not control.

5.1.3.1. Intransitive verbs

In §4.2.2.2. and §4.2.2.3., it was demonstrated that the only sociolinguistic variable that played a role in auxiliary alternation with intransitive verbs in the interview data was the socioprofessional status of speakers. In order to assess whether this social variable also has an impact on the acceptability rate of *avoir* sentences in the grammaticality judgement task, the speakers were sorted into their respective SPS level and their acceptability rates were averaged. The results can be found in Table 5.7, compared to the rates obtained in the interview.

SPS level	Average acceptability rate (in %) of the speakers	% *avoir* selection in interviews
LOW	87	14,7
MID	74	9,3
HIGH	71	8,4

Table 5.7 Average acceptability rate of intransitive sentences with *avoir* (in %) by SPS level

Table 5.7 shows that the willingness to accept intransitive verbs with *avoir* decreases as the SPS level of the speakers increases, which strongly correlates with the actual *avoir*-selection rate of the speakers, where the higher the speakers stand in the socioprofessional hierarchy the less they tend to use auxiliary *avoir* with intransitive verbs. Acceptability of auxiliary *avoir* and its actual use by speakers therefore seem to be dictated by the same social forces and with the same ranking of constraints.

Even though age was not identified as a significant factor group conditioning *avoir* usage by *Shiny Rbrul* (Johnson 2017) in the production data, in order to test whether age might impact the acceptability rates of *avoir* in the grammaticality judgement task, the speakers were sorted into their respective age categories and the acceptability rates were averaged for each, as shown in Table 5.8.

Age category	Average acceptability rate (in %) of the speakers	% *avoir* selection in interviews
Old (36-65 years old)	80	10,2
Young (18-35 years old)	75	6,6

Table 5.8 Average acceptability rate of intransitive sentences with *avoir* (in %) by age category

Results presented in Table 5.8 indicate that the age of the speakers might play a very small play a role in their readiness to accept intransitive sentences with *avoir* because the average rates of acceptability of the two age categories are fairly close to each other (75% vs 80%): older speakers accepting *avoir* variants slightly more readily than younger ones.

5.1.3.2. Pronominal verbs

In §4.3.2., distributional analyses showed that two sociolinguistic variables are at play in auxiliary alternation with pronominal verbs in the interview data: the age of the speakers and their socioprofessional level. Again, acceptability rates were averaged by age category. The results can be found in Table 5.9, compared to the actual *avoir*-selection rates obtained in the interview.

Age category	**Average acceptability rate (in %) of the speakers**	**% *avoir* selection in interviews**
Older (36-65 years old)	86	2,4
Younger (18-35 years old)	64	0,3

Table 5.9 Average acceptability rate of pronominal sentences with *avoir* (in %) by age category

Table 5.9 shows that the willingness to accept pronominal verbs with *avoir* increases alongside the age of the speaker, from 64% to 86%. This tendency was also observed in the actual *avoir*-selection rate with pronominals verbs of the speakers in the interview data: the 'old' speakers produced eight times more *avoir* tokens with pronominal verbs than 'young' ones.

Finally, the acceptability rates of each speaker were averaged based on their SPS level. Table 5.10 shows these results, compared to the actual *avoir*-selection rates with pronominals of the speakers obtained in the interviews.

SPS level	**Average acceptability rate (in %) of the speakers**	**% *avoir* selection in interviews**
LOW	94	3,5
MID	73	0,3
HIGH	59	0,2

Table 5.10 Average acceptability rate of pronominal sentences with *avoir* (in %) by SPS level

Once again, Table 5.10 shows that the willingness to accept pronominal verbs with *avoir* decreases as the SPS level of the speakers increases. This tendency was also recorded in the actual *avoir*-selection rate of the speakers during the interviews, where the higher the speakers stand in the socioprofessional hierarchy the less they tend to use auxiliary *avoir* with pronominal verbs. Equivalent to what was observed for intransitive verbs, acceptability of auxiliary *avoir* with pronominals and its actual use by speakers are dictated by the same social forces, the age and the SPS level of the speakers, and follow the same ranking of constraints.

In summary, the analysis of grammaticality judgement data of the 48 speakers by verb and by speaker showed that my speakers were much more accepting of *avoir* variants with both intransitive (with the exception of *aller, (re)venir,* and *arriver*) and pronominal verbs than what could be inferred from their production data. Although there does not appear to be a strong correlation when the judgements are examined by verb or by speaker, there are quite clear trends that emerge when they are considered according to social categories: the correlations with social variables such as the SPS and age of the speakers, while not testable statistically, are very suggestive of a real effect, since both the pronominals and non-pronominal verbs show similar patterns that also match up perfectly with the production data of the same speakers.

5.2. The crowdsourcing platform *Français de nos régions* (Avanzi *et al.* 2016)

In order to verify whether the same social predictors that I identify in my production data (see §4.2.2. and §4.3.2.) are also at play in additional Montréal French data collected by Mathieu Avanzi and André Thibault through a crowdsourcing survey in 2017, I here analyse with mixed-effects models the grammaticality judgement/self-reporting data on the two pairs of

sentences given in (80a-b). The survey asked whether the participants would say each sentence with *être,* with *avoir,* or with both indiscriminately.

(80a) *Je suis/J'ai monté sur le toit de la maison*
'I have climbed on the roof of the house'

(80b) *Je me suis/J'm'ai lavé les mains*
'I have washed my hands'

As stated in §3.6.2., after excluding certain outliers from the dataset, I obtained 808 judgements for the intransitive sentence and 805 for the pronominal one. Moreover, the tokens of the participants who were not from the Island of Montréal but from its surburbs were all collapsed into one factor group in order to investigate whether there would be geographical differences between the centre and the periphery of the Greater Montréal area.

5.2.1. *J'ai monté sur le toit de la maison* vs *Je suis monté sur le toit de la maison*

Out of the 808 answers, only 132 participants indicated that the intransitive sentence was correct with either auxiliary and 131 preferred the variant with *avoir* for an overall acceptability rate of 32,5% for *J'ai monté sur le toit de la maison*. Remarkably, this percentage is very similar to the one recorded for the *avoir* selection with the verb *(re)monter* in my sociolinguistic interviews (39,1%). Table 5.11 shows the results of the multivariate statistical analysis that was conducted on the choice of sentences in (80a) in order to determine which social factor groups play a role in the choice of variant.

MODEL 6	**Multivariate analysis of the likelihood of *J'ai monté sur le toit de la maison* (vs *Je suis monté sur le toit de la maison*)**		
Model formula: Variant.collapsed ~ City.collapsed + Education + Gender			
Input probability	0.439		
Total rate	32,5%		
Total N	808		
AIC	1000.827		
R^2	0.0462		
Deviance	990.827		
Significant factor groups	**Factor weights**	**% *avoir***	**Total N**
Level of education**	p = 0.00109		
Secondary school	0.635	56,2	32
CÉGEP (or equivalent)	0.491	41,1	163
University	0.373	29	613
Gender**	p = 0.0034		
Men	0.561	41,1	219
Women	0.439	29,4	589
City*	p = 0.0322		
Greater Montréal (excl. the Island of Montréal)	0.541	36,4	365
Island of Montréal	0.459	29,3	443
Non-significant factor groups City.collapsed:Gender City.collapsed:Education Education:Gender Age			

Table 5.11 *Shiny Rbrul* (Johnson 2017) results for Model 6: statistical effects of the significant factor groups on the likelihood of *J'ai monté sur le toit de la maison* (vs *Je suis monté sur le toit de la maison*), by *p* value and factor weights

Table 5.11 shows that the level of education is the factor group with the strongest effect, with a highly significant *p* value of 0.00109: participants who had only completed secondary school (or professional school) yielded a FW of 0.635, favouring *avoir*. The participants who had stopped their education after CÉGEP had an equal probability of selecting either sentence since their FW of 0.491 is very close to 0.5. The university-educated participants tended to disfavour the *avoir* variant, with a FW of 0.373. This result confirms my findings outlined in

§4.2.2.2. and §4.2.2.3., namely that the level of education (which was included in my calculation of the socioprofessional status) has a highly significant effect on the choice of the auxiliary with intransitive *É*-verbs. This result also confirms my findings outlined in §5.1.3.1., namely that my speakers' willingness to accept intransitive verbs with *avoir* decreases as the SPS level of the speakers increases.

While the SPS level was the only social factor group identified by *Shiny Rbrul* (Johnson 2017) as having a statistically significant effect on my 2016 data, the gender of the *Français de nos régions* participants is selected here as having a significant effect and is the factor group with the second strongest effect, with a highly significant *p* value of 0.0034. Men tended to slightly favour the sentence with auxiliary *avoir,* with a FW of 0.561, whereas women slightly disfavoured it, with a FW of 0.439. A similar but non-significant trend was observed in my 2016 production data, where the men had an *avoir*-selection rate of 11,8% and the women 9,7%.

The last factor group identified by *Shiny Rbrul* (Johnson 2017) is the municipality of the participants, with a just-significant *p* value of 0.0322. The 365 participants from the suburbs of the Greater Montréal area favoured the *avoir* variant, with a FW of 0.541, whereas the participants who live on the Island of Montréal (443 participants) slightly disfavoured *avoir,* with a FW of 0.459. Since the place of living of my speakers was included in my calculation of the SPS index, this finding is consistent with results presented in §4.2.2.2. and §4.2.2.3: it was demonstrated that my LOW SPS speakers were the most likely to use *avoir* and municipalities located on the North and South shores of Montréal were linked to a LOW SPS (see §3.2.2.4.). The above finding also parallels results outlined in §4.2.2.2.1, where it was shown that the place where the 2016 speakers spent their childhood potentially had a small effect [203] on their auxiliation patterns (see Table 4.3 above). The speakers who came from

[203] The reader will recall that the factor group 'Place of childhood' was identified by *Shiny Rbrul* (Johnson 2017) as a potentially significant factor group in Model 2 with a *p* value of 0.0458, but when added to the mixed-effects model it could no longer converge.

outside the Greater Montréal area tended to favour *avoir* (FW of 0.581) more in comparison to the ones who had grown up in the Greater Montréal area (FW of 0.419). This could lead me to believe that there is a difference in *avoir* usage depending on whether the speakers come from an urban centre or from a peripheral area.

5.2.2. *J'm'ai lavé les mains* vs *Je me suis lavé les mains*

Out of the 805 answers, 45 participants indicated that the pronominal sentence was correct with either auxiliary and 26 preferred the variant with *avoir* for an overall small acceptability rate of 8,8% for *J'm'ai lavé les mains*, four times less acceptable than the intransitive sentence with *avoir, J'ai monté sur le toit de la maison*. This percentage is almost seven times higher than that recorded for the *avoir* selection with pronominals in the sociolinguistic interviews (1,3%). While it is likely that this discrepancy could be due to a lexical effect, out of four tokens of *se laver,* none were recorded with *avoir* in the sociolinguistic interviews. Table 5.12 shows the results of the multivariate analysis that was conducted on the pairs of sentences in (80b) in order to determine which social predictors play a role in the choice of variant.

MODEL 7	**Multivariate analysis of the likelihood of *J'm'ai lavé les mains* (vs *Je me suis lavé les mains*)**		
Model formula: Variant.collapsed ~ Education + Gender			
Input probability	0.149		
Total rate	8,8%		
Total N	805		
AIC	461.914		
R^2	0.086		
Deviance	453.914		
Significant factor groups	**Factor weights**	**% *avoir***	**Total N**
Gender* **	p = 8.38e-04		
Men	0.608	15,1	218
Women	0.392	6,5	587
Level of education**	p = 0.00154		
Secondary school	0.699	28,1	32
CÉGEP (or equivalent)	0.468	12,3	162
University	0.328	6,9	611
Non-significant factor groups City.collapsed Education:Gender Age			

Table 5.12 *Shiny Rbrul* (Johnson 2017) results for Model 7: statistical effects of the significant factor groups on the likelihood of *J'm'ai lavé les mains* (vs *Je me suis lavé les mains*), by *p* value and factor weights

In the analysis of the pronominal sentence, only two factor groups were identified by *Shiny Rbrul* (Johnson 2017) as having significant effects, as shown in Table 5.12. As with the intransitive sentence, the level of education of the participants and their gender are both factor groups that condition *avoir* usage and the constraint ranking of factor levels within each of these two factor groups is identical. Here gender has the strongest effect, with a very highly significant *p* value of 8.38e-04. Men tended to favour the sentence with auxiliary *avoir* (with a FW of 0.608), whereas women disfavoured it (with a FW of 0.392). In my 2016 pronominal production data, the gender of the speakers was not believed to have an effect on auxiliary choice (see Table 4.25 above).

Table 5.12 also shows that the level of education is a strong factor group, with a highly significant *p* value of 0.00154, with participants who only completed secondary school (or professional school) yielding a FW of 0.699, favouring *avoir*. Once again, the participants who stopped their education after CÉGEP had an almost equal probability of selecting either sentences since their FW of 0.468 is very close to 0.5. The university-educated participants tended to disfavour the *avoir* variant, with a FW of 0.328. This result again is consistent with the findings outlined in §4.2.2., namely that the level of education (which was included in my calculation of the socioprofessional status) is highly significant in the choice of the auxiliary with pronominal verbs. Again, this result is also consistent with my findings outlined in §5.1.3.2., namely that my speakers' willingness to accept pronominal verbs with *avoir* decreases as the SPS level of the speakers increases.

As opposed to what was observed with the intransitive sentence *J'ai monté sur le toit de la maison* (vs *Je suis monté sur le toit de la maison*), there was no statistically significant geographical effect identified by *Shiny Rbrul* (Johnson 2017) for the *avoir* variant of the pronominal sentence, even though the distribution of the data reveals a very small effect that goes in the same direction as the one recorded for the intransitive sentence: 7,5% (33/441) of *avoir* in the pronominal sentence on the Island of Montréal and 10,4% (38/364) in the Greater Montréal (excl. the Island of Montréal). Moreover, §4.3.2. outlined that in my 2016 pronominal production data the age of the speaker was likely to be a potentially relevant variable to consider, with older speakers being more prone to favour auxiliary *avoir* than younger ones, and results from §5.1.3.2. also confirmed that the willingness to accept pronominal verbs with *avoir* increases alongside the age of the speaker. However, the age of the *Français de nos régions* participants was not identified as having a significant effect on auxiliary choice with the pronominal sentence *J'm'ai lavé les mains* (vs *Je me suis lavé les mains*).

5.3. Summary of results

In conclusion, while there are multiple variables that I was not able to control for,[204] the analysis of the grammaticality judgements data reveals that my 48 speakers were much more willing to accept *avoir* variants – with pronominal verbs even more so than with intransitive ones – than what could have been inferred from their own actual auxiliation patterns. These data yield a ranking of *avoir* acceptability for the intransitive verbs under study which confirms results outlined in §4.2.3.2., namely that there is a lexical effect, with *arriver, (re)venir,* and *aller* rarely being accepted with *avoir,* and in §4.2.5.2.1., namely the semantic effect whereby *rentrer (dedans)* and *sortir (avec)* are much more readily accepted with *avoir* than the core meanings of *rentrer* and *sortir*. With pronominal verbs specifically, the vast majority of speakers accepted most *avoir* variants and in much stronger proportions than the near-categorical *Ê*-intransitive verbs *arriver, (re)venir,* and *aller,* even though the pronominal verbs surfaced with *avoir* in the interviews only in 1,3% of the tokens. Moreover, a closer examination of the speakers' social characteristics, such as SPS and age, suggests that they are indicative of a real effect since both the intransitive and pronominal verbs showed similar patterns that also coincided with the production data of the same speakers: this confirms results presented in Chapter 4, namely that speakers belonging to the lowest SPS level favour the *avoir* variants with intransitive verbs more than the others, and that the older the speakers and the lower they are ranked in the SPS hierarchy the more easily they accept auxiliary *avoir* with pronominal sentences. The fact that the speakers from the lowest SPS accepted more *avoir* variants than the higher SPS ones could indicate that, to a certain extent, speakers ended up self-reporting their usages in their grammaticality judgement answers.

Future studies on the acceptability of *avoir* conducted through grammaticality judgement tasks should include self-reporting judgements from the speakers for all verbs in

[204] Such as semantic content, appropriateness of the register, grammaticality of another part of the sentence, and co-occurrence of other non-standard features (see §3.6.1.).

order to assess the difference in how they speak and how they think they speak. In this study, the primary goal was to obtain sufficient grammaticality judgements data in order to generate a ranking of acceptability of verbs when they surface with *avoir,* and there was a risk of not being able to collect enough data if the speakers were asked to self-report their use. Yet, during the task many speakers offered additional typical comments: "I, personally, would not say this, but you can definitely hear that around here", or "Unfortunately, some people talk like this", or even "Only country bumpkins would say something like that", etc. An analysis of these metalinguistic comments would potentially reveal interesting insights regarding linguistic attitudes and, specifically, linguistic insecurity experienced by the speakers (see §6.3.2.1.).

While the proportions of acceptable *avoir* were much higher in my grammaticality judgement tests than in the self-reporting judgements (32,5% for the intransitive sentence and 8,8% for the pronominal sentence) due to methodological differences between the two types of data collection, the recorded effects were the same across the board: the intransitive *avoir* sentences were much more readily accepted than the pronominal *avoir* ones. With data extracted from the corpus *Français de nos régions* (Avanzi *et al.* 2016), the statistical analyses presented above and produced with the help of *Shiny Rbrul* (Johnson 2017) show that, with the intransitive sentence *J'ai monté sur le toit de la maison* vs *Je suis monté sur le toit de la maison*, the level of education of the participants and their gender conditioned *avoir* use: participants who only completed secondary school and male participants were more likely to self-report the use of the *avoir* variant. It was also discovered that a geographical effect likewise accounts for the different rates of reported *avoir* use: participants who come from the Greater Montréal area (excl. the Island of Montréal) reported using the *avoir* variant significantly more than the participants who are living on the Island of Montréal. With the pronominal sentence *J'm'ai lavé les mains* vs *Je me suis lavé les mains*, the main predictor for the selection of the *avoir* variant

was the gender of the participants (male), followed by their level of education (participants who only completed secondary school).

Chapter 6. Discussion: *Le 'retour' d'un* être *cher?*[205]

I first review the evidence that has emerged from the literature review as well as from the previous two analysis chapters, and next determine how Montréal French positions itself in the broader Romance context in terms of auxiliary selection. I then explore the social implications of the major decline of *avoir* use recorded in spoken Montréal French since 1971 and, finally, assess where research on auxiliary alternation in Montréal French could most profitably be taken next.

6.1. Montréal French in the Romance context

From careful examination of the auxiliary data extracted from 48 sociolinguistic interviews, it is obvious that auxiliary selection in (Montréal) French is not at all as straightforward as prescriptivists and many syntacticians would maintain. In reference to Canadian French, both Laurentian and Acadian, Ledgeway (2019: 376) goes so far as to call its auxiliation a case of social, otherwise "free"[206], variation. As stated in §2.6., it seemed highly improbable that only one Romance variety, i.e. 'Canadian French', would show sociolinguistic variation. And in fact, variationist studies (see §2.4.2.) and multivariate analyses conducted on the 2016 Montréal French data, shown in §4.2.2., have demonstrated that auxiliation in North American French is not free at all, but conditioned by numerous social and linguistic factors.

Studies presented in §2.2.1. stated that syntactic parameters such as mood and tense (see §2.2.1.1.1. and §2.2.1.1.2.), person (see §2.2.1.1.3.), as well as argument structure (see §2.2.1.2.) dictated auxiliary selection in Romance languages according to a predictive hierarchy (see Figure 2.1). However, this type of 'categorical' one-form-per-cell distribution was actually

[205] 'The return of a loved one/*être*'
[206] See introduction of §2.2. and footnote 37 for a discussion of Romance varieties showing free variation.

not replicated with the 2016 Montréal French data because both auxiliary verbs, *être* and *avoir*, surfaced in each mood/tense, and in each subject person (with the exception of pronominal verbs, where auxiliary alternation was only recorded in 1SG). Nevertheless, some linguistic contexts did favour *avoir* more than others. For example, the past conditional tense/mood tended to favour *avoir* (as opposed to the *passé composé* which did not favour one auxiliary verb over the other or the *plus-que-parfait* which disfavoured *avoir*). With regard to the influence of person subjects, even though this variable was not selected as significantly impacting auxiliary choice in intransitive verbs, auxiliary alternation in pronominal verbs was only recorded in 1SG, recalling the "triple auxiliation" patterns described by Loporcaro (2007) in §2.2.1.1.3. where variation sensitive to person in reflexive forms is interpreted as an intermediate step towards the loss of *2-aux* in favour of a single auxiliary. However, since *avoir* rates have decreased significantly since 1971, it seems rather more likely that higher rates of *avoir* use in pronominals would have been recorded in previous corpora in comparison to 2016 rates. With respect to argument structure, Ledgeway (2019: 376) and Manente (2008: 42-43) claimed that unaccusative verbs of change of location (specifically *arriver, partir, entrer,* and *tomber*) selected *avoir* in Québécois French when they did not have a resultative meaning (see §2.2.1.2.), however my results show that these four verbs behave very differently in non-resultative contexts, since they yielded respective *avoir*-selection rates of 0,6%, 2,7%, 21,3%,[207] and 31,8% (see Table 4.9).

Moreover, the fact that there is extensive inter- and intraspeaker variation and that some verbs vary greatly in terms of their *avoir*-selection rates depending on their different meanings (literal vs figurative use) demonstrates that there is no possible alignment of my data with the Auxiliary Selection Hierarchy (Sorace 2000), which only recognises variation in *apparaître, monter* and *descendre* (§2.2.2.). While diachronic studies show that auxiliary alternation

[207] The rate given here is for verb *(re-)rentrer*, as *entrer* is systematically replaced by *rentrer* in Montréal French (see §3.2.3.1. and §3.4.).

has been recorded at least as early as Old French (see §2.1.1.), this syntactically and semantically-oriented approach also does not consider how much variation there was before the implementation of codification processes (see §2.1.2.).

The approaches presented in §2.2.1. and §2.2.2. also leave little place for sociolinguistic considerations (intra- and interspeaker variation in terms of age, gender, socioeconomic background, for example). But it does seem more likely that a variationist methodology, which typically makes use of large corpora of 'real language in use' (i.e. data collected via sociolinguistic interviews instead of via elicitation tasks) and focusses its attention on linguistic as well as social predictors by gathering extensive metadata information about the speakers themselves, would reveal social variation in the auxiliation of the Romance varieties examined in these studies.[208] Future Romance dialectological studies should therefore take into consideration more social factors because the variationist approach has shown that one-form-per-cell distributions are rare, that social factors are almost always at play, and that numerous other linguistic factors influence auxiliary alternation. So while semantic and syntactic variables can help us partially make sense of standard French split auxiliary selection, they are not sufficient to explain the variation observed all over the French-speaking world, and in the Montréal French 2016 data specifically (see §2.4.2.).

In terms of diachronic change, this study has shown change over real time, comparing data from 1971 to 2016, but the decline of *avoir* use recorded in Montréal French appears to have evolved in the opposite direction to the functionally and morphologically 'transparent' trend displayed by many Romance varieties towards a single auxiliary, namely HAVE, in the periphrastic tenses of active verbs (Lightfoot 1979: 121; Smith 1989: 311). It is noteworthy that while there has been a general trend in many Romance languages to remove auxiliary BE (see introduction to §2.2. and §2.2.1.1.3.), an extensive standardization process started in early

[208] As acknowledged in footnote 95, Ledgeway did consider sociolinguistic factors in his 2000 study and differentiates between urban, peripheral and literary Neapolitan.

modern France approximately when Canadian French and Hexagonal French split off. Supplementary diachronic evidence is required to allow us to elucidate whether much more frequent *être* use in 2016 than in 1971 signals the emergence of auxiliary *être* rather than the 'return' of *être*.[209] Especially since a 'change' to *avoir* never seems to have affected the high frequency verbs.[210] For non-frequent verbs, where most of the variation is recorded (see §4.2.5.2.2.), it might therefore be the case that extensive use of *être* was never part of 'early' Laurentian French and that in the later stages of development of this variety conservative behaviour would actually be *avoir* retention.

In the context of linguistic change in Québécois/Montréal French, the reversal of *avoir* generalization is seemingly at odds with other (morpho)syntactic changes that have not been reversed by pressure of the standard, such as the deletion of negative particle *ne* which in spoken Québécois/Montréal French has been ahead of Metropolitan French (Sankoff & Vincent 1977; Ashby 1981: 675; Martineau & Mougeon 2003), as exemplified in (81), or the increase in conditional morphology in the *si*-clauses of hypothetical complexes (Leblanc 2010), as exemplified in (82).

(81a) ***J'aime pas*** *la randonnée pédestre* (Qcois/Montréal Fr.)
'I don't like hiking'

(81b) ***Je (n')aime pas*** *la randonnée pédestre* (Metropolitan Fr.)
'I don't like hiking'

(82) ***Si je serais tombée/j'aurais tombé****, je me serais fait mal.* (Qcois/Montréal Fr.)
'If I would have fallen, I would have hurt myself'

[209] As stated in §3.2.3.16., since subordinate clauses are thought to exhibit more conservative behaviour (Lightfoot 1991: ch. 3; Harris & Campbell 1995: 27), it was hoped that the factor group 'Clause structure' would be identified by *Shiny Rbrul* (Johnson 2017) as a linguistic variable conditioning the use of *avoir* to allow me to determine whether we are witnessing an emergence of *être* rather than a 'return' of auxiliary *être*: more *avoir* use in main clauses than in embedded ones would have plausibly pointed towards a 'return' of *être*, whereas more *avoir* use in embedded clauses would have pointed towards the emergence of *être*.

[210] For the non-frequent intransitive verbs, there is also a small possibility that their auxiliation is and has always been inherently variable, even though certain (socio)linguistic variables do condition some of their auxiliation.

Sankoff & Vincent (1977, cited in Ashby 1981: 675) observed near-categorical loss of *ne* in Montréal French in 1971, and the few instances of its retention were strongly correlated with stylistic register and subject matter, but generally not linked with speakers' social characteristics. What reason could therefore explain why auxiliary levelling, proceeding in the 'organic' historical trajectory observable across many Romance languages, is being reversed by the standard in Montréal French, unlike, say, *ne* deletion and *si* + conditional, which are both explicitly proscribed and corrected in school? Apart from the fact that a 'wrong' auxiliary might be more (phonologically) salient, the answer may lie in the social distribution of the variable under study.

6.2. The social distribution of auxiliary *avoir*

6.2.1. With intransitive verbs

In the 2016 Montréal French data, the multivariate analyses presented in §4.2.2.2. and §4.2.2.3., as well as the distributional analyses presented in §4.2.3., §4.2.4., and §4.2.5., do indeed reveal that an analysis combining a multiplicity of both semantic and syntactic variables must also incorporate social ones in order to fully account for how the variation in auxiliation is distributed.[211]

While it is true that my sampling methodology might have resulted in a slight magnification of the overall changes since 1971 (see §3.2.2.4.), and therefore that *avoir* rates in 2016 might actually be slightly higher than stated, one key element is undeniable: auxiliary

[211] With intransitive verbs, the semantic variables conditioning auxiliary alternation are '(In)animacy of subject', 'Hodiernal action', and 'Verb meanings'. The syntactic variables are the 'Presence of a pronominal clitic before the auxiliary' and the 'Presence of intervening material between the auxiliary and the past participle', as well as the 'Possibility of transitive use'. The 'Tense of the verb' is the only syntax-semantics variable that conditions *avoir* use and the 'SPS' is the only social variable that conditions *avoir* selection.

alternation in intransitive verbs has significantly decreased in Montréal French in 45 years (from 32,8% in 1971 to 10,7% in 2016), but a small variation persists mostly in the speech of LOW SPS speakers (see Tables 4.2 and 4.5 above). The fact that new statistical tools, such as mixed-effects models, applied on newly collected data yield different results from the previous studies (Sankoff & Thibault 1977; Thibault & Sankoff 1997) is not entirely surprising, but comparisons of my findings with those of Sankoff (2019), who also employed mixed-effects models, show that most linguistic variables that condition *avoir* use in the 2016 data also had significant effects on the combined 1971, 1984, and 1995 corpora (§4.2.6.). My study does confirm many of Sankoff's recent observations regarding her earlier data, but also introduces a new variable tested for the first time on French auxiliary alternation data and which happened to have a very highly significant effect on *avoir* selection, namely, whether a pronominal (non-reflexive) clitic appears before the auxiliary. The presence of such an element before the auxiliary was one of the strongest predictor of *avoir* use across mixed-effect models. It turned out to be the strongest predictor of *avoir* use when tested on the most variable lexical items (presented in Model 3 in Table 4.5) and the second strongest predictor in Model 2 (presented in Table 4.2), which included all the verbs tested by Sankoff & Thibault (1977).

One must therefore look more closely at social dynamics such as increased access to education (see §3.2.2.4.) as well as exposure to the standard language (see §2.4.2.2.4.) in order to explain the real-time observation that auxiliary *être* is now more present in the speech of francophone Montréalers than in 1971. Even though dialect contact did not seem to have an impact on the 2016 auxiliary data, it might be the case that immigration from France has played a role in the decrease of *avoir*. Data published on the 8th of April 2020 by the Consulat général de France à Québec,[212] show that in the last 10 years, the (Metropolitan) French

[212] https://quebec.consulfrance.org/La-communaute-francaise-au-Quebec [webpage accessed on May 13, 2020]

community in the province of Québec has grown at a very high rate: the number of French citizens established in Québec has practically doubled since 2005, going from 45,890 to 80,900. Every year, between 3,000 and 4,000 French citizens permanently move to Québec: as of December 31, 2019, close to 100,000 (99,289) French citizens were registered as living in Canada, out of which 61,550 were based in Montréal.[213]

As stated in §1.1. and in §2.4.2.2.2., since Sankoff (2019: 220) recorded a lower *avoir*-selection rate in 1984 than in 1971, it led her to conclude that the patterning of the data across age groups represented a change in progress towards the diminution of *avoir*: her real-time results confirm her apparent-time interpretation. By contrast, there was no significant difference according to date of recording, suggesting that speakers did not change as they aged and that her results were therefore not indicative of age grading (2019: 208). Sankoff consequently concludes that it was an instance of "lifespan stability in the face of community linguistic change" (2019: 211):

> Adult speakers are progressively being replaced by younger community members who enter the picture with different grammars. What is being transmitted to children? I would suggest that children are keen observers of vectors within the speech community, and that by the time they are adolescents, they are well aware of the fact that it is old-fashioned to say, as grandfather does, *j'ai venu au monde* 'I came into the world [was born]'. (2019: 221)

Since no age effect was identified with my data, it could mean that we are witnessing the final stages of the change, yet a small stable variation would seem to persist in the speech of lower SPS speakers (see §4.2.2.2. and §4.2.2.3.). Since the SPS has continued to have a significant effect, this suggests that the nature of change between 1971-1984-1995 (Sankoff 2019: 220) and 2016 is still a change from above (the level of consciousness), where the change is adopted first by higher status and educated speakers (Labov 2006) because they have greater access to standard norms.

[213] See https://quebec.consulfrance.org/La-communaute-francaise-au-Quebec for additional details. [webpage accessed on May 13, 2020]

With regard to change through space, the evidence gathered in §4.2.2.2.1. and §5.2.1. suggests that speakers from the Montréal urban centre are less frequent users of auxiliary *avoir* compared to those from more rural areas surrounding Montréal. Trudgill's 'gravity model' (1974), which predicts that linguistic innovations radiate from larger, more densely populated centres to smaller and peripheral ones, might suggest that a more frequent use of *être* would indeed be a Montréal innovation and that, in time, it is likely to reach the peripheries (of Québec).

6.2.2. With pronominal verbs

With pronominal verbs, the interpretation must be rather speculative because of the few variable data gathered. The distributional analysis presented in Table 4.25 (§4.3.2.) shows that auxiliary alternation with pronominal verbs is very restricted, with only six speakers out of 48 using *avoir*. The person of the subject is the best predictor of *avoir* selection since 100% of *avoir* pronominal tokens analysed surfaced with 1SG subjects. In descriptive terms, the variation is also socially conditioned because among the six speakers that conjugated a pronominal verb with *avoir,* four belonged to the lowest SPS, one to the middle one, and one to the highest one. Additionally, the five older variable speakers selected *avoir* three times more frequently than the remaining younger one. The fact that the *avoir* tokens were mostly uttered by older speakers suggests that the use of *avoir* with pronominal verbs is likely to be even more restricted in the future, probably to the oldest speakers of the lowest SPS. As to gender, the four female speakers produced an *avoir* rate twice as high (13/42, 31%) as the two male ones (3/21, 14,3%). This is unexpected, since female speakers are expected to lead in linguistic change from above the level of consciousness (see (§2.4.2.2.1. above for a brief explanation of that part of the Gender Paradox), but this result could be explained by the fact that out of the four women, three were LOW SPS speakers and one MID SPS, while only one man was

a LOW SPS speaker and the other, a HIGH SPS one. The gender result cannot be explained by an interaction with age since the only 'young' speaker to use *avoir* with a pronominal verb was a woman.

With regard to linguistic conditioning, the presence of a clitic object pronoun before the auxiliary categorically triggers *avoir* (100%) for these speakers as opposed to when there is no clitic pronoun (19%). The tokens that had an intervening element between the auxiliary and the past participle display an *avoir* rate of 37,5%, whereas those which did not surfaced with *avoir* with a lower rate (23,6%). As for the influence of tense, the tokens in *conditionnel passé* selected *avoir* with a 100% rate, the *plus-que-parfait* ones with a 42,9% rate, and the *passé composé* ones with a 18,9% rate (see §4.3.2.). It is remarkable that these main linguistic effects are also statistically significant with the intransitive data, therefore suggesting that the auxiliation of pronominal verbs behaves in very similar patterns to the intransitive ones, albeit with more restricted conditioning (i.e. with the added influence of the speakers' age and the person of the subject). However, the major difference in *avoir* rates between intransitive verbs and pronominals verbs most probably pertains to the social meaning of the variable, which I address below.

6.2.3. An indicator, a marker, or a stereotype?

Indicators are dialectal variables that only display variation on the social or geographical level (i.e. among different social classes or different geographical areas) but not on the formality continuum (Labov 1972b: 314-317). The difference between markers and stereotypes resides in the level of consciousness: speakers are aware of stereotypical linguistic features, which can then be subject to metapragmatic discussion, but speakers are not aware of ones that are markers (Labov 1972b: 314-317).

As mentioned in the introduction, it was in an Oxford classroom that I first became aware of the existence of auxiliary alternation in French (with the 1971 Montréal French data). If left me truly astonished because, as a native speaker, I was completely oblivious to the fact that it had been/was part of the variety of French that my fellow Montréalers spoke. By conducting this present research, I aimed to discover whether my reaction then was caused by the fact that the variation was now almost extinct or rather by the fact that it was not salient at all for speakers. I did not expect to find so many instances of *avoir* use overall in my corpus, but can confirm that my intuitions (see §1.1.) regarding the saliency of certain marked forms were accurate. While I do not currently have enough stylistic variation data to determine whether auxiliary alternation in every intransitive verb is an indicator, marker or stereotype, it is clear that the uses of *tomber* and of pronominal verbs with *avoir* are stereotypes. Normally, it is claimed that stereotypes are salient to both in-group and out-group members, but with this particular variable it could also be the case that for speakers who use auxiliary *avoir* with pronominal verbs this specific variant is an indicator or marker, but for the non-users it behaves as a stereotypical feature.

As the previous chapter has shown, the analysis of the grammaticality judgement data reveals that the 48 speakers were much more willing to accept *avoir* variants than could have been inferred from their own actual auxiliation patterns (see §5.1.). That was the case with pronominal verbs even more so than with intransitive ones. Verbs *arriver, (re)venir,* and *aller* are rarely accepted with *avoir,* and *rentrer (dedans)* and *sortir (avec)* are much more readily accepted with *avoir* than the core meanings of *rentrer* and *sortir*. Speakers belonging to the lowest SPS level favour the *avoir* variants with intransitive verbs more than the others. With the data from the corpus *Français de nos régions* (Avanzi *et al.* 2016), the statistical analyses presented above show that with the intransitive sentence *J'ai monté sur le toit de la maison* vs *Je suis monté sur le toit de la maison* the level of education of the participants and their gender were highly significant

variables, participants who had only completed secondary school and male participants being predictors of higher levels of self-reporting as using the *avoir* variant (see §5.2.1.).

Evidence to suggest that most speakers are aware that the surfacing of auxiliary *avoir* with the verb *tomber* is (negatively) connoted can be found in how this variant is employed by parodic content creators for humoristic purposes, as evidenced by the two figures reproducing humoristic content by using *tomber* and pronominal verbs with *avoir* (Figures 6.1 and 6.2).[214] For example, the online Montréal magazine *Urbania* staged a fake humourous Facebook Messenger chat conversation between Canada's current Prime Minister Justin Trudeau and Steven Guilbeault, former director of *Équiterre*[215] and *Greenpeace Québec* as well as current Minister of Canadian Heritage. During this exchange, Trudeau tries to convince Guilbeault to accompany him to the inauguration ceremony of the new Samuel-De Champlain Bridge. The quality of Canadian PM Justin Trudeau's spoken French has long been criticized (in the media[216]), and the content creator behind this parody capitalizes on this well-known fact in order to mock Trudeau, who at the time was in the hot seat because of the *Trans Mountain* pipeline saga. A still frame from the video is presented in Figure 6.1.

[214] I was not able to find any other instance of humoristic use of auxiliary *avoir* with any other intransitive verb.
[215] *Équiterre* is a Canadian non-profit and non-governmental organization, operating in Québec. It manages a community-supported agriculture system of farms and consumers, including households and institutions. (https://www.equiterre.org/en)
[216] See in particular Bosworth (2019), and newspaper pieces by Patch-Neveu (2019, in *Le Devoir*) and Brousseau-Pouliot (2017, in *La Presse*) criticizing Trudeau's "English" syntax when he speaks (and sometimes writes in) French. It is often said that he speaks in *bilingue* 'bilingual'.

Figure 6.1 Screen capture (1m17s) of a fake humoristic Facebook chat between current Canadian PM Justin Trudeau and former director of *Équiterre* and *Greenpeace Québec*/current Minister of Canadian Heritage Steven Guilbeault (taken from the Facebook page of Montréal-based online magazine *Urbania* and dated June 26, 2019)[217]

Justin Trudeau's last chat message from the still frame is reproduced and translated in (83). It contains an instance of *tomber* conjugated with auxiliary *avoir* in the past infinitive.

(83) *J'ai pensé à ça après* ***avoir tombé*** *sur mes vieux cahiers de notes de mon cours d'écologie avec les 3R: réuser, réducer, récycler.*
'I thought about it after stumbling across my old notebooks from my ecology class with the three Rs: reuse, reduce, recycle'

The comedian behind this exchange wants to draw attention to the ungrammaticality of Trudeau's French by combining this 'wrong' *avoir* variant in addition to three invented verbs

[217] https://www.facebook.com/watch/?v=581370695724711 [webpage accessed on May 13, 2020]

réuser, réducer, and *récycler* (the first two verbs being modelled on their English equivalents "reuse" and "reduce"), rather than using the French verbs *réutiliser, réduire,* and *recycler.*[218]

As the previous chapter has shown, speakers seem to be aware of the existence of pronominal verbs conjugated with *avoir* in the local variety of French even when they do not themselves use that variant (see §5.1.2.2.). Moreover, the vast majority of speakers accepted most pronominal *avoir* variants (§5.1.1.4.) and in much greater proportions than the near-categorical *Ê*-intransitive verbs *arriver, (re)venir,* and *aller,* even though the pronominal verbs surfaced with *avoir* during the interviews almost as rarely (only in 1,3% of cases) as these near-categorical intransitive *Ê*-verbs (see §5.1.1.2.). The older the speakers and the lower they are ranked in the SPS hierarchy the more easily they accept auxiliary *avoir* with pronominal sentences. With the pronominal sentence *J'm'ai lavé les mains* vs *Je me suis lavé les mains* from *Français de nos régions* (Avanzi *et al.* 2016), the main predictor for the self-reporting selection of the *avoir* variant was also the gender of the participants (male), followed by their level of education (participants who had only completed secondary school) (see §5.2.2.).

There is additional evidence to suggest that for most speakers the surfacing of auxiliary *avoir* with pronominal verbs is negatively connotated by being socially marked. Once again, evidence can be found in how this variant is appropriated in humoristic content. Figure 6.2 shows a meme from the parodic Québécois Instagram account *SkedooSled* (@skedoosled), which mocks the 'white trash'/crass lifestyle of young working-class Québecers living in rural areas and stereotypically obsessed with snowmobiles. The text is reproduced with a translation in (84).

[218] It is unclear however whether the author of this parody used the variant *avoir tombé* in order to signal once again the influence of English structure on Trudeau's speech patterns or to simply mock his 'bad' French.

Figure 6.2 Screen capture of an Instagram post[219] published by the Québécois meme account *SkedooSled* (dated February 28, 2020)

(84) *Le ptit Maxence gelé ben tight qui se décaliss drette dans l'poteau après* ***s'avoir tiré*** *un gros criss de plomb de hash*
'Lil' Maxence, super high, who crashes straight into the (electricity) pole after having smoked[220] a big fucking hash bottle toke/hit'

From these two illustrations, it is obvious that the use of *avoir* with *tomber* and with a pronominal verb are stereotypes (at least for their audience) because, in being used for humoristic effect, they are subject to metapragmatic discussion. I do not know whether such uses of *avoir* would have prompted the same comedic effect on the 1971 population or on current users of *avoir* with pronominals. It is moreover interesting to note that both examples occur with a past infinitive, suggesting that the *avoir* variant is most salient to speakers in this non-finite context.

[219] https://www.instagram.com/p/B9HkTBchVV0/ [webpage accessed on May 13, 2020]
[220] In European varieties of French, the expression *se tirer un plomb* rather means 'to shoot oneself'.

6.3. Directions for future research on auxiliary alternation

This study suggests that further work on auxiliary alternation should incorporate inputs from two different axes. This section will examine which additional data sources it would be profitable to use to further triangulate my results. I then explore the potential contribution of other linguistic fields and analytical tools on auxiliary alternation research and the kind of information their use might reveal.

6.3.1. Comparisons with other data sources

My results could be first compared to older Laurentian French corpora, then with Montréal neighbourhoods corpora, and lastly to corpora of other Canadian French varieties, in order to find the answers to some research questions regarding auxiliary alternation that remain unsolved to this day.

6.3.1.1. Comparison with earlier Laurentian French data

Since the oldest Montréal French data only goes back to the 1960s,[221] a comparison with even earlier corpora of Laurentian French could possibly help us determine whether high frequency verbs have ever been conjugated with *avoir* or in higher proportions. This might tell us whether we are witnessing a 'return' of auxiliary *être*, rather than the progressive loss of auxiliary *avoir,*

[221] In 1962-1963 at the Université du Québec à Montréal (UQAM), André Dugas and Gilles Bibeau (Université de Montréal) conducted a series of interviews with 102 participants, out of which 71 were either born in Montréal or had moved there before the age of 5 (Dugas 1986). Their corpus included 45 men and 57 women from 14 different neighbourhoods (information found in Boisvert & Laurendeau 1988: 243-245). The corpus is not balanced in terms of age and occupation, as very few speakers were over 50 years old and 48 women out of the 57 indicated that they were housewives. The interviews were either semi-structured or unstructured, but in general the interview themes covered the biography of the participant: personal history, work, family, sports, hobbies, cultural activities. Each recording lasted about 30 minutes and was roughly transcribed by *une habile secrétaire de souche québécoise* 'a skilled secretary of Québécois origin' (information found in Boisvert & Laurendeau 1988: 243-245). Out of the 102 recordings, 71 were subjected to a normalized transcription that respects syntactic order but does not take into account pronunciation features and restores for example pronouns, articles, conjunctions omitted in speech, except in the contexts of verb morphology, where the form originally pronounced is indicated (quoted in Boisvert & Laurendeau 1988: 243-245). To my knowledge, data from this corpus have very rarely been used as comparison points for subsequent linguistic studies on Montréal French.

or whether these verbs have never been subject to auxiliary variation (see §2.6.). One possible dataset could be Shana Poplack's *Récits du français québécois d'autrefois* (2009), a compendium of folk tales, legends, and personal interviews of 44 native francophones born between 1846 and 1895, including speakers from five (administrative) regions of Québec[222] but not Montréal. These audio materials (approximately 90 hours) were recorded between 1942 and 1955, have been standardized, and rendered machine readable. With the same line of reasoning, it might be revealing to examine early modern French data stemming from the geographical areas that provided settlers for Nouvelle-France in order to see whether there was auxiliary alternation, and in what conditions it occurred. To this end, France Martineau's *Corpus de français familier ancien* (Corpus LFFA, 2005) would be a good starting point for exploration since it comprises family correspondence (more than 15 000 letters) in French vernacular, divided over diatopic (French America and Northwest France), diachronic (17th, 18th, and 19th centuries), and diastratic axes (lower and upper social classes).

An exhaustive verification of the state of auxiliary alternation with pronominal verbs in the 1971-1984-1995 Montréal French corpora might also prove useful in order to determine whether the minor auxiliary alternation observed in 2016 indicates a stable variation or one that is decreasing with time.

6.3.1.2. Comparison with Montréal neighbourhoods data

Since it proved difficult to find native speakers of Montréal French (see §3.1.1.), the question that remains to be asked is whether that label in linguistic research is even relevant or useful anymore. It is most likely the case that 'Montréal French' is now an umbrella term that encompasses an array of varieties, since Montréal is a growing and increasingly diverse

[222] The five regions are Gaspésie-Îles-de-la-Madeleine [Gaspésie, Baie des Chaleurs], Capitale-Nationale [Québec, Charlevoix, Malbaie], Chaudière-Appalaches [Beauce], Saguenay-Lac-St-Jean, and Côte-Nord, some of which are known to have furnished input settlers to the francophone community of Ottawa-Hull. See http://www.sociolinguistics.uottawa.ca/holdings/canadian-fe.html for more information on the corpus.

metropolis. Since I have shown that the geographical definition of 'Montréal' has changed over the years (see §3.1.1.), it is expected that the typological definition of 'Montréal French' has evolved alongside it. In line with this reasoning, linguists such as Blondeau *et al.* (2012, 2013) and Bigot & Papen (2018) have decided to compile corpora of individual Montréal neighbourhoods, mentioned in §3.1.1., Hochelaga-Maisonneuve/Saint-Michel-Montréal-Nord, and Ahuntsic-Cartierville, respectively.[223]

The Saint-Michel-Montréal-Nord (2013) variationist corpus also aims to document the contribution to the dynamics of spoken Montréal French by young Montréalers of immigrant background. There is evidence that both Haitian creole and Arabic are two strong influences on the language of young Montréalers (Guidara 2018). It would therefore be very unrepresentative and exclusionary to refrain from considering these second- or third-generation immigrants as native speakers of Montréal French. It would also be reductive to keep on studying native 'old-stock' Montréal French speakers, when there are regular arrivals in Québec of French native speakers from all-over the Francophonie, who undoubtedly have an influence on local speakers.

A study of auxiliary alternation in these neighbourhood corpora, especially in the speech of young participants, would be key to triangulate my findings and verify whether the same rates of *avoir* as well as the same social and linguistic conditioning can be observed with these speakers.

[223] There is also an older Montréal French corpus that focusses on the Centre-Sud neighbourhood. Doran, Drapeau, & Lefebvre (1982) at UQAM conducted 43 interviews between 1976 and 1978 in Montréal, within a quadrilateral delimited by the streets Sherbrooke, Sainte-Catherine, Amherst, and De Lorimier. The 43 participants, 18 men and 25 women, are all adolescents and pre-adolescents between the ages of 6 and 18 years-old (information found in Boisvert & Laurendeau 1988: 250-251). In order to have a control group of parents, six adults aged between 42 and 52 years old were also interviewed. The adolescents and pre-adolescents were recorded in different 'natural' communication situations with their friends, their parents, during various activities: games, group activities, etc. The interviews with the parents were the only ones that were structured. The corpus contains approximately 40 hours of recording (information found in Boisvert & Laurendeau 1988: 250-251).

6.3.1.3. Comparison with other Canadian French varieties

In order to verify whether the decrease in auxiliary alternation observed in Montréal French follows a path that is similar across other Canadian French varieties, it would be important to triangulate my findings with recent corpus data from other Canadian French varieties. This could tell us whether different stages of a development can be observed or whether this variable follows different paths of development, if at all, and exhibits different social and linguistic conditioning according to different speech communities.

It would also be quite relevant to explore the self-reports/grammaticality judgement data of other Canadian and European French speakers compiled by the *Français de nos régions* (Avanzi *et al.* 2016) corpus to verify how they differ from the Montréal data.

6.3.2. Additional subfields of sociolinguistic study

Throughout this research, it was mentioned that further insights might be gained from other fields of linguistic study, such as linguistic attitudes, accommodation theory, and stylistic variation.

6.3.2.1. Linguistic attitudes

Sankoff (2019: 221) mentions that attitudinal factors should be taken into consideration when studying language change across the lifespan, and that this also holds true for language change across a community. It is for this reason that at the end of each sociolinguistic interview I asked my participants to fill in a questionnaire in French about linguistic attitudes (see Appendix 3D), as mentioned in Chapter 3. Based on Labov's findings in Martha's Vineyard (1963), this questionnaire aimed at observing the potential role of identity in auxiliary alternation. The questionnaire includes questions such as:

1. Do you feel more Canadian, Québécois, or Montréaler?
2. Is it important that we continue to speak French in Québec?
3. What is your opinion of Bill 101? [224]
4. Which variety of French do you value more, the standard or *franco-québécois*?
5. What do you think of the state/quality of Québec French now?
6. What do you think of the way French is taught in schools, and is used in the media?
7. Do you think it is important to speak French/English with Québec Anglophones?
8. What do you think of Québécois music, films, literature?
9. How do you think could Montréal French be improved?

In order to analyse the results of the linguistic attitudes questionnaire, a future study could attempt to verify whether it is possible to correlate pro-Québec and pro-Québécois French sentiments with speakers who have a stronger *avoir* usage. This would potentially indicate that *avoir* generalization may be a defining characteristic of Québécois French: the equivalent of the "local" variant in Labov's Martha's Vineyard study (1963).

6.3.2.2. Accommodation theory and priming

The basis of communication accommodation theory, which was developed by Giles & Smith (1979) among others, relies on the main principle that people adjust their style of speech to one another: speakers adopt convergent and divergent behaviours of communication, "particularly as they relate to people's goals for social approval, communication efficiency, and identity" (Ebesu Hubbard 2009: 121). With this in mind, a closer look at the 2016 data might reveal that an *avoir* token is more likely to follow another *avoir* token, when uttered by an interlocutor (the interviewer or another speaker).

[224] See §2.6. and footnote 97 for a brief explanation of its controversial aspects.

It might even be the case that an *avoir* token is more likely to be primed by the presence of a previous one when both are uttered by the same speaker, for example if two *Ê*-verbs are juxtaposed in a sentence and appear one right after another.

6.3.2.3. Style-shifting/ attention paid to speech

No variationist study on auxiliary alternation had examined the potential impact of intraspeaker stylistic variation and, in agreement with Sankoff (2019: 219), I believe that new fieldwork focused on eliciting data from the same speaker according to different formality registers or to how much attention is being paid to speech (see §4.2.5.1.1.) might reveal important information regarding the status of the variable in the community (whether it is an indicator or marker). In addition, insight provided by the anecdote in §3.1.3.[225] suggests that it would be interesting to check whether we could observe in the speech of some participants higher rates of false starts with *avoir* followed by self-corrections with *être* in more formal registers, indicating that for such speakers the variable is above the level of consciousness.

[225] After a technical glitch in my recording equipment, speaker Johanne and her sister ended up answering my interview questions a second time and used used far fewer *avoir* tokens the second time around. See footnote 117.

Chapter 7. Conclusion

Before the start of this project, the only occurrences of auxiliary *avoir* use of which I was personally aware were with *tomber* (and mostly in the 1SG) and with pronominals,[226] as exemplified in (85a-b)(see §1.1. and §6.2.3.). Hence the title of my thesis, which has investigated the auxiliary alternation in spoken Montréal French between *avoir* and *être* with the twenty-or-so intransitive verbs prescriptively requiring the latter (*tomber*, *partir, rester,* etc.), as well as with pronominal verbs, which had been left out of most previous studies on auxiliary alternation.

(85a) *J'ai tombé* (AVOIR) vs *Je suis tombé* (ÊTRE)
'I fell/have fallen'

(85b) *Je m'ai fait mal* (AVOIR) vs *Je me suis fait mal* (ÊTRE)
'I (have) hurt myself'

After analysing the Sankoff-Cedergren Montréal Corpus (1971), Sankoff & Thibault (1977: 107) observed a rate of *avoir* of 34% and concluded that greater exposure to the standard slowed it down. Using a trend study – by creating a new sociolinguistic corpus of Montréal French[227] and by analysing the speech of 48 native speakers of Montréal French during semi-structured interviews – I have examined in detail the state of this alternation today given the great sociodemographic changes that have taken place in Montréal in the last 50 years (see §3.1.1.). Since 1971, the province of Québec has witnessed an important wave of nationalism that strongly correlated with the designation of French as its only official language, with the promulgation in 1977 of the Charter of the French language.

[226] As stated in §1.1., I believed them to be very rare or highly socially marked.
[227] Once fully transcribed, my corpus will be made available for consultation by other linguists.

Chapter 2 established that many different factors can affect auxiliary selection in French as well as in other Romance languages. On the one hand proponents of semantically-oriented syntactic approaches (Sorace 2000, 2004; Legendre & Sorace 2003; Legendre 2007, 2017; Giancarli 2011, 2015) have mostly concentrated their efforts on studying the auxiliation patterns of standard French (see §2.2.2.), while on the other hand some Romance syntacticians and dialectologists (Ledgeway 2019; Loporcaro 2007, 2016) have been able to make comparative observations by collecting a limited amount of data but from an extensive range of non-standard Romance dialects (see §2.2.1.). Both approaches focus on just a few linguistic environments (mood, tense, subject person, unaccusativity, semantic features, etc.) and leave very little space for the investigation of inter- and intraspeaker variation as well as social variables, even though numerous studies of auxiliary selection in North American French dialects have shown that such speaker variations exist and that social factors are always at play (see §2.4.2.); and specifically in spoken Montréal French. In fact, spoken varieties of French have largely been omitted from (semantic-)syntactic analyses, except in rare cases where there is mention of Sankoff & Thibault's Montréal study (1977).

The research goals of my project (see §2.6.) were therefore to determine the state of auxiliary alternation in Montréal French today compared to 1971 (as well as to 1984 and 1995), and in which direction the changes had been observed (see §4.2.1.) I also wanted to examine whether there had been a change in the social and linguistic distribution and conditioning of the variable and, if so, what form it had taken (see §4.2.2.), as well as whether these intransitive and pronominal data would turn out to be distributed in similar social and linguistic patterns to those that had already been recorded in other French and Romance varieties (see §2.2.). Subsequently, another goal of the project was to determine what would be the implications of such findings (see Chapter 6). I also wanted to examine the state of auxiliary alternation in pronominal verbs in Montréal French (see §4.3.) and test whether it was possible to confirm my overall findings through two triangulation methods (see Chapter 5): the analysis of

grammaticality judgement data and self-reporting judgements from the survey *Français de nos régions* (Avanzi *et al.* 2016).

In addition to recording a substantial decrease – by more than two-thirds – of *avoir*-selection rates in intransitive verbs since 1971, a mixed-effects modelling variationist methodology (see Chapter 3) conducted with statistical software *Shiny Rbrul* (Johnson 2017) revealed that various linguistic factors and a few sociolinguistic ones that had not been previously uncovered – by neither Romance (semantic-)syntactic theorists (see §2.2.), nor by previous sociolinguistic studies on French (see §2.4.2.) – do condition *avoir* selection in Montréal French. This is notably the case of the inanimacy of the subject, the presence of a clitic pronoun before the auxiliary verb, a pre-hodiernal action, and the place of childhood of the speakers (see §4.2.2.2. and §4.2.2.3.).

As to my second and third reseach goals, a direct comparison of linguistic and social effects with older data was not possible without using outdated statistical tools, but Sankoff's timely updated work of 2019 allowed me to confirm that practically the same linguistic independent variables which were at play in the previous corpora of Montréal French and they were also significant in my 2016 data. It also allowed me to confirm that the age effect which was present in 1971, 1984, and 1995 is no longer observable in 2016, suggesting that the small residual variation will remain stable, with more frequent *avoir* tokens surfacing in the speech of lower SPS speakers and in the linguistic contexts that favour *avoir* (see §4.2.2.2. and §4.2.2.3.). My main results indicate that the two most significant predictors of *avoir* selection in intransitive verbs are the inanimacy of the subject and the presence of a clitic pronoun appearing between the subject and the auxiliary verb. The variable '(In)animacy of the subject' had never been tested with Montréal French data[228] and the variable 'Presence of a pronoun

[228] At the time of testing, Sankoff (2019) had not yet been published.

between the subject and the auxiliary' had never been tested in any study on auxiliary alternation in French (see Model 2 in Table 4.2 and Model 3 in Table 4.5 above).

Another original contribution of this thesis is the exhaustive study of auxiliary alternation in pronominal verbs in Montréal French (see §4.3.). I have shown that while the variation is mostly restricted to older SPS speakers, similar linguistic factors are also at play, such as the 'Presence of a pronoun between the subject and the auxiliary' and 'Presence of an intervening element between the auxiliary and the past participle'. But unlike with intransitive verbs, the inanimacy of the subject produces the opposite effect with pronominals since 100% of *avoir* tokens were triggered by 1SG subjects and therefore were animate ones (see §4.3.2.).

An additional unique contribution lies in the extensive triangulation of my data with two additional corpora: 48 grammaticality judgements collected after every sociolinguistic interview (see §5.1.) and self-reporting judgements on two pairs of sentences (*J'ai monté sur le toit de la maison* vs *Je suis monté sur le toit de la maison* and *J'm'ai lavé les mains* vs *Je me suis lavé les mains*) collected by the crowdsourcing platform *Français de nos régions* (Avanzi *et al.* 2016) in 2017 (see §5.2.). The comparative analysis of grammaticality judgements revealed that speakers belonging to the lowest SPS level favour the *avoir* variants with intransitive verbs more than the others, and that the older the speakers and the lower they are ranked in the SPS hierarchy the more easily they accept auxiliary *avoir* with pronominal sentences. While not tested statistically, these correlations with social variables such as the SPS and age of the speakers were suggestive of a real effect since both the pronominal and non-pronominal verbs showed patterns that were not only similar but also entirely matched up with the production data collected during the interviews (see §5.1.3.). The grammaticality judgement task also revealed that the 48 speakers seemed to be aware of the existence of pronominal verbs conjugated with *avoir* in the local variety of French even when they do not themselves make use of that variant (see §5.1.1.). With self-reporting data from *Français de nos régions* (Avanzi *et al.* 2016), the main predictor for the selection of the *avoir* variant was the gender of the

participants (male), followed by their level of education (participants who only completed secondary school). A small geographical effect was also discovered with the intransitive *avoir* variant: participants who come from the Greater Montréal area reported using the *avoir* variant significantly more than the participants who lived on the Island of Montréal (see §5.2.1.).

While I still do not know whether we are witnessing the return or the emergence of *être* in this variety of French, the decline of *avoir* use in Montréal French appears to illustrate a gradual (re)alignment with standard French or a process of *dévernacularisation* (Mougeon 2005), and thus a reversal of its long trajectory towards a single auxiliary system, in the opposite direction of what has taken place in the history of multiple Romance languages (such as standard Castilian, Portuguese, Aromanian, standard Catalan, Sicilian, many Ibero-Romance and many far southern Italian varieties, see §2.2.). As stated in §2.4.2.2.6., Bigot (2011: 13) had observed that "les membres des élites sociale et culturelle du Québec emploient de façon homogène un modèle grammatical oral très proche de celui présenté dans *Le bon usage* (donc de l'écrit)."[229] This (re)alignment of spoken Québécois French with the (European) standard is furthermore confirmed by Molinari & Puccini (2013, non-paginated) who regard it as one of the results of the stigmatization of Québécois French:

> La stigmatisation dont le français québécois a fait l'objet s'est souvent traduite en une stigmatisation culturelle et identitaire et, par conséquent, en une insécurité dont les Québécois ont essayé de sortir en suivant deux voies principales : d'une part, la revendication de leur propre spécificité identitaire, culturelle et linguistique ; de l'autre l'alignement sur le français hexagonal, variété valorisée à tous les niveaux.[230]

[229] 'the members of Québec's social and cultural elites use in an homogenous way a spoken grammatical model very close to the one outlined in *Le bon usage* (and therefore to the written language)'

[230] 'The stigmatization to which Québécois French has been subjected was often translated into a cultural and identity stigmatization and, consequently, into an insecurity from which the Quebecers tried to extricate themselves by adopting two main solutions: by claiming a distinct identity and their own cultural and linguistic specificity on the one hand, and, on the other hand, by aligning their French with Metropolitan French, the variety that is valued at all levels.'

In order to fight against the linguistic insecurity widespread in Québec, Sankoff & Cedergren tried to emphasize with the creation of their corpus in 1971 the systematic character of Québécois French linguistic features that were traditionally denigrated, as mentioned in the introduction and in §2.4.2.1. Their work also gave special consideration to the metalinguistic discourse of participants and, seeking inspiration from this objective, I created in 2016, concomitantly with this research, a linguistic attitudes questionnaire that was completed by my 48 participants at the end of the sociolinguistic interview (see §6.3.2.1. and Appendix 3D). The data from these questionnaires have not yet been exploited, but future research will be able to make use of these valuable data in order to provide a portrait of current linguistic representations occurring in Montréal. Up-to-date works investigating such matters are sparse, even though the demolinguistic profile of Québec has been greatly transformed in the last decades, while linguistic insecurity is still rampant in Québec and far from innocuous in its consequences.[231]

[231] See, for example, Remysen (2018, 2020) on the consequences of linguistic insecurity, particularly on Québécois children and teenagers.

References

Ackema, P. & A. Sorace (2017). 'Auxiliary Selection', In M. Everaert & H. van Riemsdijk *The Wiley Blackwell Companion to Syntax*, 2nd ed. John Wiley & Sons, 1-32.

Allard, M. (2013). 'Il y a 70 ans, l'école devenait obligatoire', *Le Devoir* [Montréal], May 25. Web. https://www.ledevoir.com/opinion/idees/379028/il-y-a-70-ans-l-ecole-devenait-obligatoire

Andersen, H. (1990). 'The structure of drift", In H. Andersen & K. Koerner (eds.) *Historical Linguistics 1987: papers from the 8th International Conference on Historical Linguistics*, Lille, 31 August-4 September 1987, Amsterdam & Philadelphia: John Benjamins, 1-20 (*Current Issues in Linguistic Theory* 66).

Andriani, L. (2017). *The syntax of the dialect of Bari*. PhD thesis: University of Cambridge.

Antonini, A. (1753). *Principes de la grammaire francoise, pratique et raisonnée*, Paris: Duschesne.

Aranovich, R. (2003). 'The Semantics of Auxiliary Selection in Old Spanish', *Studies in Language* 27: 1-37.

Aranovich, R. (2007). 'Split auxiliary selection from a cross-linguistic perspective', In R. Aranovich (ed) *Split Auxiliary Systems: A cross-linguistic perspective,* Amsterdam: John Benjamins.

Armstrong, N. (2001). *Social and Stylistic Variation in Spoken French: A comparative approach*, Amsterdam: John Benjamins.

Ash, S. (2013). 'Social Class', In J.K Chambers, P. Trudgill & N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*, 2nd ed., Oxford: Blackwell, 350-367.

Ashby, W. J. (1981).'The loss of the negative particle *ne* in French: a syntactic change in progress', *Language* 57: 674-687.

Auger, J. (2003). 'The development of a literary standard: the case of Picard in Vimeu-Ponthieu, France', In B. D. Joseph et al (ed.) *When Languages Collide*, Columbus: The Ohio State University Press, 141-164.

Avanzi, M., C. Barbet, J. Glikman & J. Peuvergne (2016). 'Présentation d'une enquête pour l'étude des régionalismes du français', *Online proceedings of the 5th Congrès Mondial de Linguistique Française*, Tours (France), July, not paginated. https://doi.org/10.1051/shsconf/20162703001

Avanzi, M. & A. Thibault (2018). 'Réflexions épistémologiques sur de nouveaux apports méthodologiques et empiriques à l'étude géolinguistique des français d'Amérique', *Online proceedings of the 6th Congrès Mondial de Linguistique Française*, Mons (Belgium), July, not paginated. https://doi.org/10.1051/shsconf/20184602001

Avram, L. (1994). 'Auxiliary configurations in English and Romanian', *Revue roumaine de linguistique* 5-6: 493-510.

Avram, L. & V. Hill (2007). 'An irrealis BE auxiliary in Romanian', In R. Aranovich (ed.) *Split Auxiliary Systems. A Cross-linguistic Perspective*, Amsterdam/Philadelphia: Benjamins, 47-64.

Bailey, G., T. Wikle, J. Tillery & L. Sand (1991). 'The apparent time construct', *Language Variation and Change* 3: 241-264.

Balcom, P. (2008). 'On the learning of auxiliary use in the referential French variety by speakers of New Brunswick Acadian French', *Canadian Journal of Linguistics/Revue canadienne de linguistique* 53(1): 7-34.

Bard Gurman, E, Robertson, D. & A. Sorace (1996). 'Magnitude Estimation of Linguistic Acceptability', *Language* 72 (1): 32-68.

Barrière, C. (2009). 'L'anglicisation continue', *Le Droit* [Gatineau], May 12. Web. https://www.ledroit.com/archives/langlicisation-saccentue-14b271ced7cd1b0ea0946bb836bbe209

Batlle, M. (2002). *L'expressió dels temps compostos en la veu mitjana i la passive pronominal: El procés de substitució de l'auxiliar 'ésser' per 'haver'*. Barcelona: Institut d'Estudis Catalans/ Publicacions de l'Abadia de Montserrat.

Battye, A, M.-A. Hintze & P. Rowlett (2000). *The French language today: A linguistic Introduction*, 2nd ed., London: Routledge.

Bauche, H. (1920). *Le Langage populaire. Grammaire, syntaxe et dictionnaire du français tel qu'on le parle dans le peuple avec tous les termes d'argot usuel*, Paris: Payot.

Bauche, H. (1946). *Le Langage populaire. Grammaire, syntaxe et dictionnaire du français tel qu'on le parle dans le peuple avec tous les termes d'argot usuel*, 4th ed., Paris: Payot.

Bayley, R. (2013). 'The Quantitative Paradigm', In J.K Chambers & N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*, 2nd ed., Oxford: Blackwell, 85-107.

Bayley, R., R. Cameron & C. Lucas (2013). *The Oxford Handbook of Sociolinguistics*, Oxford: Oxford University Press.

Beauchemin, N., P. Martel & M. Théoret (1992). *Dictionnaire de fréquence des mots du français parlé au Québec,* New York: Peter Lang.

Beaudoin-Bégin, A.-M. (2015). *La langue rapaillée : combattre l'insécurité linguistique des Québécois*, Montréal: Somme toute.

Beaulieu, L. (1995). *The social function of linguistic variation: A sociolinguistic study in a fishing community of the north-eastern coast of New Brunswick*. PhD thesis: University of South Carolina.

Beniak, E. & R. Mougeon (1989). 'Recherches sociolinguistiques sur la variabilité en français ontarien', In E. Beniak & R. Mougeon (eds.) *Le français canadien parlé hors Québec: Aperçu sociolinguistique,* Québec: Les Presses de l'Université Laval, 69-104.

Bentley, D. & T. Eythórsson (2001). 'Alternation According to Person in Italo-Romance', in L. Brinton (ed.) *Historical Linguistics 1999: Selected Papers from the 14th International Conference on Historical Linguistics*. Amsterdam/Philadelphia: Benjamins, 63-74.

Bentley, D. & T. Eythórsson (2003). 'Auxiliary selection and the semantics of unaccusativity', *Lingua* 114: 447-71.

Benveniste, E. (1965). 'Structure des relations d'auxiliarité', *Acta Linguistica Hungariae*, IX, 1, 1-15.

Bernard Barbeau, G. (2017). 'Disqualification d'autrui, disqualification de soi : l'auto-dévaluation chez les Québécois francophones', *Minorités linguistiques et société* 8: 83-101.

Bertinetto, P. M. & M. Squartini (2016). 'Tense and aspect', In A. Ledgeway & M. Maiden (eds.) *The Oxford Guide to the Romance Languages*, Oxford: Oxford University Press, 939-953.

Bescherelle, L.-N. (2006). *L'art de conjuguer : dictionnaire de 12 000 verbes*, revised by C. Contant, Montréal: Hurtubise.

Bigot, D & R. Papen (2018). 'Deux nouveaux corpus de français parlé au Québec : Sainte-Anne-des-Lacs et Montréal (Ahuntsic-Cartierville)', *Les français d'ici 2018*, Université Concordia, Montréal, May 23-25.

Bigot, D. (2011). 'De la norme grammaticale du français parlé au Québec', *Arborescences* 1 (March): Identités linguistiques, langues identitaires : à la croisée du prescriptivisme et du patriotisme. https://doi.org/10.7202/1001939ar

Blanc, M. (1994). 'French in Canada', In C. Sanders (ed.) *French today: language in its social context French language today*, Cambridge: Cambridge University Press.

Blanche-Benveniste, C. (1977). 'L'un chasse l'autre, le domaine des auxiliaires', *Recherches sur le Français* [*sic*] *parlé,* n° 1: 100-148.

Blanche-Benveniste, C. (1997). 'La notion de variation syntaxique dans la langue parlée', *Langue française,* n°115: 19-29.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Boberg, C. (2004). 'Real and apparent time in language change: Late adoption of changes in Montreal English', *American Speech* 1, 79(3): 250-269.

Boisvert, L. & Laurendeau, P. (1988). 'Répertoire des corpus québécois de langue orale', *Revue québécoise de linguistique* 17 (2): 241-261. https://doi.org/10.7202/602634ar

Bosworth, Y. (2019). 'The "Bad" French of Justin Trudeau: When Language, Ideology, and Politics Collide', *American Review of Canadian Studies*, 49(1): 5-24. https://doi.org/10.1080/02722011.2019.1570954

Bourdieu, P. (1991). *Language and symbolic power*, Cambridge, MA: Harvard University Press.

Bourdieu, P. & L. Boltanski (1975). 'Le fétichisme de la langue', *Actes de la recherche en sciences sociales* 4, Paris, 2-32.

Boychuk, J. (2017). 'Special Report Living English: Where are Quebec's Anglos? Map of population distribution from 2011 Census', *Canadian Broadcasting Corporation* [Montréal]. Web. http://www.cbc.ca/montreal/features/livingenglish/anglophone-census-map/[Accessed April 11, 2017]

Brasseur, P. & J.-P. Chauveau (1990). *Dictionnaire des régionalismes de Saint-Pierre-et-Miquelon*, *Canadiana Romanica* 5, Tübingen: Max Niemeyer Verlag.

Brasseur, P. (2000). 'La mise en œuvre du corpus dans le traitement des données lexicographiques : l'exemple du franco-terre-neuvien', In C. Bavoux, R. Dupuis, J.-M. Kasbarian (eds.) *Images de la variation, perception et modélisation de la variation du français, en hommage à Daniel Baggioni*, Paris: L'Harmattan, 193-216.

Brousseau-Pouliot, V. (2017). 'Notes biographiques de Justin Trudeau : 5 fautes de français en 560 mots', *La Presse*, May 31.

Brown, B. (1988). 'Problems with past participles [*sic*] agreement in French and Italian dialects', In D. Birdsong & J.-P. Montreuil (eds.) *Advances in Romance Linguistics,* Dordrecht: Foris Publications, 51-66.

Buchard, A. & A. Carlier (2008). 'La forme verbale « *être* + participe passé » en tant que marqueur d'aspect et de la structure argumentale : une typologie graduée', In J. Durand, B. Habert, B. Laks, (eds.) *Congrès Mondial de Linguistique Française – CMLF'08*, Paris: Institut de Linguistique Française, 2421-2437.

Buridant, C. (2000). *Grammaire nouvelle de l'ancien français.* Paris: Sedes.

Burzio, L. (1986). *Italian Syntax: A Government-Binding Approach*, Dordrecht: Reidel.

Bybee, J. (2010). 'Markedness: iconicity, economy and frequency', In J. Jung Song (ed.) *The Oxford Handbook of Linguistic Typology*, Oxford: Oxford University Press.

Cajolet-Laganière, H., P. Martel & C.-É. Masson (with Louis Mercier as editorial advisor) (2013). *Dictionnaire Usito : Parce que le français ne s'arrête jamais*, produced in the framework of the Franqus project, Sherbrooke: Édition Delisme [Online].

Cajolet-Laganière, H. & P. Martel (1993). 'Entre le complexe d'infériorité linguistique et le désir d'affirmation des Québécois et Québécoises', In M. Francard *et al.* (ed.) *L'insécurité linguistique dans les communautés francophones périphériques*. *Actes du colloque de Louvain-la-Neuve*, November 10-12, *Cahiers de l'Institut de linguistique de Louvain* 19 (3-4): 169-185.

Calzolari, N. (1996). 'Observação e Generalização : Análise Linguística de Verbos declaratives Italianos com Base em Corpora Linguísticos', In F. Bacelar Do Nascimento, M.C. Rodrigues, & J. Bettencourt Gonçalves (eds.) *Actas do XI Encontro National da Associação Portuguesa de Linguística*, Vol. I, Corpora. Lisbon: Colibri, 93-106.

Canale, M., R. Mougeon & M. Bélanger (1977a). *Aspects of the Use of the Auxiliary AVOIR in Ontarian French*, Toronto: Centre de recherches en éducation franco-ontarienne, Institut d'études pédagogiques de l'Ontario: The Ontario Institute for Studies in Education.

Canale, M., R. Mougeon & M. Bélanger (1977b). 'Aspects de l'usage de la préposition POUR en français ontarien: interférence et/ou surgénéralisation?', *Working Papers on Bilingualism* 12: 61-78.

Canale, M., R. Mougeon & M. Bélanger. (1978). 'Analogical leveling of the auxiliary *être* in Ontarian French', In M. Suñer (ed.) *Contemporary Studies in Romance Linguistics*, Washington, D.C.: Georgetown University Press, 41-61.

Cedergren, H. (2018). Round table *À l'écoute du français d'ici, Les français d'ici 7,* Concordia University, Montréal, May 23.

Cedergren, H. & D. Sankoff (1974). 'Variable rules: performance as a statistical reflection of competence', *Language* 50(2): 333-55.

Chambers, J. K. (2003). *Sociolinguistic Theory*, 2nd ed. Oxford, UK: Blackwell.

Chambers, J.K. (2009). *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*, 3rd ed. Oxford: Wiley-Blackwell.

Chambers, J.K. (2013). 'Patterns of Variation including Change', In J.K Chambers & N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*, 2nd ed., Oxford: Blackwell, 297-323.

Chaudenson, R., Mougeon, R. & É. Beniak (1993). *Vers une approche panlectale de la variation du français,* Institut d'Études Créoles et Francophones, Didier Érudition, Coll. "Langues et développement".

Chiominto, C. (1984). *Lo parlà forte della pora ggente*. Rome: Bulzoni.

Chomsky N. (1965). *Aspects of the Theory of Syntax*, Cambridge, MA: MIT Press.

Clark, E. V. (1987). 'The principle of contrast: A constraint on language acquisition', In B. MacWhinney (ed.) *Mechanisms of Language Acquisition*, London: Lawrence Erlbaum Associates.

Clark, E. V. (1993). *The Lexicon in Acquisition*, Cambridge: Cambridge University Press.

Cocchi, G. (1995). *La selezione dell'ausiliare*. Padua: Unipress.

Connors, K. (1975). 'L'effacement de *que* - règle syntaxique', *Recherches linguistiques à Montréal* 4: 17-33.

Corriveau, J. (2016). 'Hochelaga-Maisonneuve s'embourgeoise', *Le Devoir* [Montréal], July 6. Web. http://www.ledevoir.com/politique/montreal/474877/etude-hochelaga-maisonneuve-s-embourgeoise [Accessed April 11, 2017]

Croft, W. (2003). *Typology and universals*, 2nd ed., Cambridge: Cambridge University Press.
Cukor-Avila, P. (2012). 'Some structural consequences of diffusion', *Language in Society* 41: 615-640.

Cukor-Avila, P. & G. Bailey (2011). 'The interaction of transmission and diffusion in the spread of linguistic forms', *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 39*. 17: 41-49.

Cukor-Avila, P. & G. Bailey (2013). 'Real and apparent time', In J.K Chambers & N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*, 2nd ed., Oxford: Blackwell, 239-62.

Culbertson, J. (2004). '*Gradient Auxiliary Selection: Support for a Semantic Approach to Unaccusativity*.' Honors thesis, Brown University.

D'Alessandro, R. & I. Roberts (2010). 'Past participle agreement in Abruzzese: Split auxiliary selection and the null-subject parameter', *Natural Language and Linguistic Theory* 28, 41-72.

Damourette, J. & E. Pichon (1911-1940). *Des Mots à la pensée. Essai de grammaire de la langue française*, Paris: D'Artrey.

Dauby, J. (1979). *Le livre du "rouchi". Parler picard de Valenciennes*, Amiens: Musée de Picardie.

Deriano, P. (2005). *Grammaire du gallo, langue romane de Haute-Bretagne*. Ploudalmézeau, Bretagne: Éditions Label LN.

Descusses, M. (1986). *Le Patois ardennais de Gespunsart*, Paris: Société d'études linguistiques et anthropologiques de France.

Deshaies, D. & E. LaForge (1981). 'Le Futur simple et le futur proche dans le français parlé dans la ville de Québec', *Langues et Linguistique* 7, 21-37.

Deulofeu, H.-J. (2017). 'Les constructions comme variables pour la variation syntaxique et le rôle des traditions discursives dans la production des variants', *Revue de Sémantique et Pragmatique* 41-42: 117-134.

Dion, N. (2003). *L'effacement du* que *en français canadien : Une étude en temps réel*. Master's dissertation: University of Ottawa.

Doran, M., L. Drapeau & C. Lefebvre (1982). 'Le projet Centre-Sud: corpus et méthodologie', In C. Lefebvre (ed.) *La syntaxe comparée du français standard populaire : approches formelles et fonctionnelles*, vol. 2, Office de la langue française, coll. Langues et sociétés, 401-420.

Dorian, N. (1978). 'The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic', *Language* 54: 590-609.

Dragomirescu, A. (2010). *Ergativitatea. Tipologie, sintaxă, semantică*, Bucharest: Editura Universității din București.

Dragomirescu, A. & A. Nicolae (2009). 'Relics of Auxiliary Selection in Romanian', Iorgu Iordan − Al. Rosetti Institute of Linguistics, Bucharest and University of Bucharest: unpublished manuscript.

Dugas, A. (1986). *Le corpus Bibeau-Dugas*, internal document, Université du Québec à Montréal.

Eager, C. & J. Roy (2017). 'Mixed Effects Models are Sometimes Terrible', University of Illinois at Urbana Champaign. https://arxiv.org/pdf/1701.04858.pdf

Ebesu Hubbard, A. S. (2009). 'Perspective taking, adaptation, and coordination', In W. F. Eadie (ed.) *21st Century communication : A reference handbook,* vol. 1, Los Angeles, London, New Dehil, Singapore, Washington DC: *SAGE Publications.*

Eckert, P. (2012). 'Three waves of variation study: The emergence of meaning in the study of variation', *Annual Review of Anthropology* 41: 87-100.

Emirkanian, L. & D. Sankoff (1985). 'Le futur 'simple' et le futur 'proche', In M. Lemieux & H. Cedergren (eds.) *Les Tendances dynamiques du français parlé à Montréal,* vol. 1, Québec: Office de la langue française, 189- 204.

Estienne, H. (1569), *Traité de la conformité du langage français avec le grec.*

Flutre, L.-F. (1955). *Le Parler picard de Mesnil-Martinsart (Somme)*, Geneva: Droz.

Fónagy, I. (1985). 'J'aime. Je connais. Verbes transitifs à objet latent', *Revue Romane* 20(1): 3-35.

Fontaine, J.A. (1888). 'On the History of the Auxiliary Verbs in the Romance Languages', *University Studies of the University of Nebraska* I.

Förster, A. (1908). Avoir *und* être *als Hilfsverb beim intransitiven Zeitwort in ihrer Entwicklung vom Alt- zum Neufranzösischen*. Darmstadt : Otto.

Fournier, J.-M. (2004). 'Crise de langue et conscience linguistique : la question de la règle des 24 heures', *Dix-septième siècle* 2(223): 251-264.

Frei, H. (1929). *La grammaire des fautes*, Paris: Geunther.

Gagnon, K. (2016). 'Embourgeoisement insidieux à HoMa?', *La Presse* [Montréal], July 6. Web. http://www.lapresse.ca/actualites/montreal/201607/05/01-4998312-embourgeoisement-insidieux-a-homa.php [Accessed April 11, 2017]

Gauchat, L. (1905). 'L'unité phonétique dans le patois d'une commune', *Aus Romanischen Sprachen und Literaturen: Festchrift Heinrich Mort*. Halle: Max Niemeyer, 175-232.

Gesner, E. (1978). 'L'emploi des auxiliaires *avoir* et *être* dans le parler acadien de la Baie Sainte-Marie, Nouvelle-Écosse', In G. W. Patterson (ed.) *Papers from the second annual meeting of the Atlantic provinces linguistics association*, Halifax: Mount Saint Vincent University, 16-22.

Giancarli, P.-D. (2011). *Les auxiliaires* être *et* avoir*: Étude comparée corse, français, acadien et anglais,* Rennes: Presses Université de Rennes.

Giancarli, P.-D. (2015). 'Auxiliary Selection with Intransitive and Reflexive Verbs: the limits of gradience and scalarity, followed by a proposal', In R. Kailuweit & M. Rosemeyer (eds.) *Auxiliary Selection Revisited: Gradience and Gradualness*, FRIAS book series "Linguae & Litterae", vol. 44, Berlin/Münich/Boston: De Gruyter, 79-120.

Giles, H. & P. Smith (1979). 'Accommodation Theory: Optimal Levels of Convergence', In H. Giles, & R.N. St. Clair (eds.) *Language and Social Psychology*, Baltimore: Basil Blackwell.

Godard, D. (1992). 'Le programme labovien et la variation syntaxique'. *Langages* 26(108), *Hétérogénéité et variation : Labov, un bilan*, 51-65.

Gougenheim, G. (1929). *Étude sur les périphrases verbales de la langue française*. Paris: Les Belles lettres.

Gougenheim, G. (1951). *Grammaire De La Langue Française Du Seizième Siècle*. Collection *Les Langues Du Monde. Série Grammaire, Philologie, Littérature* 7. Lyon: IAC.

Grevisse, M. & A. Goosse. (2008). *Le bon usage. Grammaire française avec des remarques sur la langue française d'aujourd'hui*, 14th ed., Gembloux: Duculot.

Grevisse, M. (1964). *Problèmes de langage*, 3rd series, Paris: Presses universitaires de France.

Grevisse, M. (1969). *Le Bon Usage*, Gembloux: Duculot.

Grevisse, M. (2011). *Le bon usage. Grammaire française refondue par André Goosse*, 15th ed., Gembloux: Duculot.

Guidara, A. (2018). 'Petit glossaire de l'argot montréalais/Le nouveau joual de la métropole', *La Presse* [Montréal], September 9.
https://plus.lapresse.ca/screens/a483b787-a330-418d-904d-4a8814866b2a-__7C___0.html

Guiraud, P. (1969). *Le Français populaire*, Paris: Presses Universitaires de France.

Haden, E. F. (1973). 'French dialect geography in North America', In T. A. Sebeok (ed.) *Current Trends in Linguistics. Vol 10: Linguistics in North America*, The Hague: Mouton, 422-444.

Hallion, S. (2000). *Étude du français parlé au Manitoba*. PhD thesis: Université de Provence, Aix-en-Provence.

Hallion Bres, S. (2006). 'Similarités morphosyntaxiques des parlers français de l'ouest canadien', In R. Papen & G. Chevalier (eds.) *Les variétés de français en Amérique du Nord. Évolution, innovation et description, Revue canadienne de linguistique appliquée-Revue de l'Université de Moncton* 37(2): 111-131.

Harris, A. & L. Campbell (1995). *Historical Syntax in Cross-Linguistic Perspective,* Cambridge: Cambridge University Press.

Hays, M. (2016). 'The one-in-six rule: can Montreal fight gentrification by banning restaurants?', *The Guardian* [London, UK], November 16. Web. https://www.theguardian.com/cities/2016/nov/16/one-in-six-rule-can-montreal-canada-fight-gentrification-banning-restaurants [Accessed April 11, 2017]

Heidinger, S. (2015). 'The auxiliary selection of French *monter* 'move upward' from the 16th to the 20th century', In R. Keiluweit & M. Rosemeyer (eds.) *Auxiliary Selection Revisited. Gradience and Gradualness*. Berlin: De Gruyter, 277-300.

Heinen, S. & H. Kadow (1990). 'The Acquisition of French by Monolingual Children: A Review of the Literature', In J. Meisel (ed.) *Two First Languages – Early Grammatical Development in Bilingual Children*. Dordrecht: Foris, 47-71.

Hendschel, L. (2012). *Li Croejhete walone: contribution à une grammaire de la langue wallonne*. Retrieved November 1, 2009, from http://users.skynet.be/lorint/croejh/node75.html.

Hermann, M. E. (1929). 'Lautveränderungen in der Individualsprache einer Mundart', *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen*, *Philosophisch-historische Klasse* 11: 195-214.

Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.

Hoffman, M. (2014). 'Sociolinguistic Interviews', In J. Holmes & K. Hazen (eds.) *Research Methods in Sociolinguistics: A Practical Guide,* John Wiley & Sons, 25-41.

Institut de la statistique du Québec (2000). 'Les régions métropolitaines', *La situation démographique au Québec, bilan 2000*, 21-36. Web. http://www.stat.gouv.qc.ca/statistiques/population-demographie/bilan00-regions-metrop.pdf [Accessed April 11, 2017]

Johnson, D. E. (2009). 'Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis', *Language and Linguistics Compass* 3(1): 359-383.

Johnson, D. E. (2010). 'Why stepwise isn't so wise', invited panel presentation at New Ways of Analyzing Variation (NWAV) 39, San Antonio.

Johnson, D. E. (2014). *Progress in regression: Why natural language data calls for mixed-effects models*. Self-published manuscript.

Johnson, D. E. (2017). 'Progress in Regression: Statistical and Practical Improvements to Rbrul', Paper presented at NWAV 46 (2-5 November), Madison (Wisconsin).

Jones, M. C. (1988). 'Ergative and impersonal constructions in Sardinian', Paper presented at the *Sixteenth Cambridge Romance Linguistics Seminar* (January 1988), Trinity Hall, Cambridge.

Jones, M. C. (2001). *Jersey Norman French. A Linguistic Study of an Obsolescent Dialect.* (Publications of the Philological Society 34.) Oxford: Blackwell Publishing.

Jones, M. C. (2008). *The Guernsey Norman French Translations of Thomas Martin: A Linguistic Study of an Unpublished Archive* (Orbis Supplementa 31; Monographs published by the International Center of General Dialectology (Louvain)). Leuven: Paris; Dudley, MA: Peeters.

Jones, M. C. (2015). *Variation and Change in Mainland and Insular Norman*. Leiden, Nederland: Brill.

Kailuweit, R. (2015). 'BE or HAVE in Contemporary Standard French – residua of semantic motivation', In R. Kailuweit & M. Rosemeyer (eds.) *Auxiliary Selection Revisited: Gradience and Gradualness*, FRIAS book series "Linguae & Litterae", vol. 44, Berlin/Münich/Boston: De Gruyter, 249-276.

King, R. & T. Nadasdi (2000). 'How Auxiliaries Be/Have in Acadian French', In P. Balcom *et al.* (eds.) *Papers from the 24th Annual Meeting of the Atlantic Provinces Linguistic Association*, Moncton: Université de Moncton, 61-72.

King, R. & T. Nadasdi (2003). 'Back to the future in Acadian French', *French Language Studies* 13, 323-337.

King, R. & T. Nadasdi (2005). 'Deux auxiliaires qui voulaient mourir en français acadien', In P. Brasseur & A. Falkert (eds.) *Français d'Amérique: Approches morphosyntaxiques,* Paris: L'Harmattan, 103-11.

Kiparsky, P. (1973). 'Productivity in phonology', In M. Kenstowitz & C. Kisseberth (eds.) *Issues in Phonological Theory*, The Hague: Mouton.

Kircher, R. (2012). 'How pluricentric is the French language? An investigation of attitudes towards Quebec French compared to European French', *French Language Studies* 22: 345-370.

Kraus, A. (2006). 'Language attitudes of Québécois students towards *le français québécois standard* and *le franco-québécois*.' Master's dissertation: University of North Carolina at Chapel Hill.

Kukenheim, L. (1967). *Grammaire historique de la langue française*. Leiden: Universitaire Pers Leiden.

Kurath, H., B. Bloch & M. L. Hansen (1939). *Linguistic atlas of New England*. Providence, R.I: Brown University.

Labelle, M. (1992). 'Change of state and valency', *Journal of Linguistics* 28: 375-414.

Laberge, S. (1977). *The changing distribution of indeterminate pronouns in discourse*. Mimeo.

Labov, W. (1963). 'The social motivation of a sound change', *Word* 19: 273 -309.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Labov, W. (1971). 'The study of language in its social context', In J. A. Fishman (ed.) *Advances in the Sociology of Language*, vol. 1., The Hague: Mouton, 152-216.

Labov, W. (1972a). *Language in the Inner City: Studies in the Black English Vernacular*, Philadelphia, PA: University of Pennsylvania Press.

Labov, W. (1972b). *Sociolinguistic Patterns*, Philadelphia, PA: University of Pennsylvania Press.

Labov, W. (1978), 'Where does the Linguistic Variable Stop ?, A Response to Beatriz Lavandera', *Working Papers in Sociolinguistics, Southwest Educational Development Laboratory*, Austin.

Labov, W. (1982). 'Building on empirical foundations', In W. P. Lehmann & Y. Melkiel (eds.) *Perspectives on historical linguistics*, Amsterdam and Philadelphia: John Benjamins, 17-92.

Labov, W. (1994). *Principles of Linguistic Change, vol. 1: Internal Factors*, Cambridge, MA: Blackwell.

Labov, W. (2001). *Principles of Linguistic Change, vol. 2: Social Factors*, Cambridge, MA: Blackwell.

Labov, W., Ash, S. & C. Boberg (2005). *The Atlas of North American English: Phonetics, Phonology and Sound Change Volume 1*, Berlin: De Gruyter.

Lacetera, M. (1982). *Zagarolo: un dialetto, una cultura, un modo di essere*. Rome: Trevi.

Lamarre, P. (2016). 'Une francophonie canadienne en redéfinition : transformations, brouillages et tensions discursives', Paper presented at *Les français d'ici*, Université de Saint-Boniface, Winnipeg, Manitoba, 7-9 June 2016.

Larjavaara, M. (2000). *Présence ou absence de l'objet. Limites du possible en français contemporain*, Helsinki: Academia Scientiarum Fennica.

Lavandera, B. R. (1978). 'Where Does the Sociolinguistic Variable Stop?', *Language in Society* 7(2): 171-182.

Léard, J.-M. (1995). *Grammaire québécoise d'aujourd'hui - Comprendre les québécismes*, Montréal: Guérin Universitaire.

Leblanc, C. (2010). 'Tracing a morphosyntactic change in Quebec French: The non-standard conditional in *si*-clauses', *Journal of French Language Studies* 20(2): 151-169.

Ledgeway, A. (2000). *A Comparative Syntax of the Dialects of Southern Italy: A Minimalist Approach*. Oxford: Blackwell.

Ledgeway, A. (2011). 'Syntactic and morphosyntactic typology and change', In M. Maiden, J. C. Smith & A. Ledgeway (eds.) *The Cambridge History of the Romance Languages: Structures,* vol. I, Cambridge: Cambridge University Press, 382-471.

Ledgeway, A. (2012). *From Latin to Romance: Morphosyntactic Typology and Change*, Oxford: Oxford University Press.

Ledgeway, A. (2014). 'Romance Auxiliary Selection in Light of Romanian Evidence', *Diachronic Variation in Romanian,* 3-34.

Ledgeway, A. (2019). 'Parameters in the development of Romance perfective auxiliary selection', In M. Cennamo & C. Fabrizio (eds.) *Historical Linguistics 2015. Selected Papers from the 22nd International Conference on Historical Linguistics*, Amsterdam: Benjamins, 343-384.

Legendre, G. (2007). 'Optimizing auxiliary selection in Romance', In R. Aranovich (ed.) *Split Auxiliary Systems: A Cross-Linguistic Perspective*, Amsterdam: John Benjamins, 145-180.

Legendre, G. (2017). 'Auxiliaries', In A. Dufter & E. Stark (eds.) *Manual of Romance Morphosyntax and Synatx*, Berlin, Boston: De Gruyter, 272-298.

Legendre, G. & A. Sorace (2003). 'Auxiliaires et intransitivité en français et dans les langues romanes', In D. Godard (ed.) *Les Langues Romanes: Problèmes de la Phrase Simple*. Paris: CNRS editions, 185-234.

Levitt, J. (1979). 'Variations in the use of the auxiliairies *avoir* and *être* in modern French', *Geolinguistics* 5, Bridgeport, CT: American Society of Geolinguistics, 25-32.

Lightfoot, D. (1991). *How to Set Parameters. Arguments from Language Change*, Cambridge: Cambridge University Press.

Loporcaro, M. (1999). 'L'ausiliazione perfettiva nelle parlate di Zagarolo e di Colonna e lo studio della sintassi dei dialetti mediani', *Contributi di filologia dell'Italia mediana* 13: 203-26.

Loporcaro, M. (2007). 'On triple auxiliation in Romance', *Linguistics* 45(1): 173-222.

Loporcaro, M. (2016). 'Auxiliary selection and participial agreement', In A. Ledgeway & M. Maiden (eds.) *The Oxford Guide to the Romance Languages*, Oxford: Oxford University Press, 802-18.

Macaulay, R. K. S. (1977). *Language, social class and education: a Glasgow study*. Edinburgh, UK: Edinburgh University Press.

Mackenzie, I. (2006). *Unaccusative Verbs in Romance Languages*. London: Palgrave Macmillan.

Maiden, M. (2015). 'Le roumain : présentation grammaticale', *Lalies* 35: 9-88.

Maiden, M. & C. Robustelli (2007). *A Reference Grammar of Modern Italian,* London: Hodder Arnold.

Manente, M. (2008). *L'aspect, les auxiliaires "être" et "avoir" et l'hypothèse inaccusative dans une perspective comparative français/italien*. Unpublished PhD thesis: Universities of Venice and Paris VIII.

Manzini, M. R. & L. Savoia (2005). *I dialetti italiani e romanci. Morfosintassi generativa* (3 vols), Alessandria: Edizioni dell'Orso.

Marceau, J. (2016). 'L'embourgeoisement dans Hochelaga, le Plateau d'il y a 30 ans', *Radio-Canada* [Montréal], June 21. Web. http://ici.radio-canada.ca/nouvelle/787916/hochelaga-maisonneuve-gentrification-embourgeoisement-echec-mixite-sociale-plateau [Accessed April 11, 2017]

Marchello-Nizia, C., B. Combettes, S. Prevost & T. Scheer (eds.) (2020). *Grande grammaire historique du français.* Boston: De Gruyter Mouton.

Martineau, F. (1985). *Élision variable de (que) dans le parler d'Ottawa-Hull.* Master's dissertation: Université d'Ottawa.

Martineau, F. (1988). 'Variable deletion of *que* in the spoken French of Ottawa-Hull', In D. Birdsong & J.-P. Montreuil (eds.) *Advances in Romance Linguistics*. Foris: Dordrecht, 275-287.

Martineau F. (dir.) (2005-). *Le Corpus LFFA (laboratoire de français familier ancien).* www.polyphonies.uottawa.ca

Martineau, F. (2018). 'Variation et variétés : Fluidité des frontières', In F. Martineau, A. Boudreau, Y. Frenette, & F. Gadet (eds.), *Francophonies nord-américaines : langues, frontières et idéologies*, coll. Les Voies du français. Québec: Presses de l'Université Laval.

Martineau, F. & R. Mougeon (2003). 'A sociolinguistic history of the origins of *ne* deletion in European and Quebec French', *Language* 79: 118-152.

Martinet, A. (1969). *Le français sans fard*, Paris: Presses Universitaires de France.

Mateu, J. (2009). 'Gradience and auxiliary selection in Old Catalan and Old Spanish', In P. Crisma & G. Longobardi (eds.) *Historical Syntax and Linguistic Theory*, Oxford : Oxford University Press.

Meyerhoff, M. (2011). *Introducing Sociolinguistics*, 2nd ed., Abingdon/New York: Routledge

Milroy, J. (1992). *Language and Social Networks,* Baltimore, MD: University Park Press.

Milroy, L. & M. Gordon (2003). *Sociolinguistics: Method and Interpretation,* Malden, MA: Blackwell.

Moignet, G. (1973). *Grammaire de l'ancien français : morphologie, syntaxe*. Paris : Klincksieck.

Molinari, C. & P. Puccini (2013). 'Autour du français québécois: perspectives (socio-) linguistiques et identitaires', *Repères DoRiF: Voix/voies excentriques: la langue française face à l'altérité*, no 2. http://www.dorif.it/ezine/ezine_printarticle.php?id=77 [Accessed April 23, 2014]

Montpetit, J. (2016). 'How Montreal's St-Henri neighbourhood has become a battleground in Canada's gentrification debate', *Canadian Broadcasting Corporation News* [Montréal], May 31. Web. http://www.cbc.ca/news/canada/montreal/gentrification-st-henri-montreal-toronto-vancouver-1.3607655 [Accessed April 11, 2017]

Motapanyane, V. (ed.) (2000). *Comparative Studies in Romanian Syntax*, Amsterdam, Elsevier.

Mougeon, R. (2005). 'Rôle des facteurs linguistiques et extra-linguistiques dans la dévernacularisation du parler des adolescents dans les communautés francophones minoritaires', In A. Valdman, J. Auger & D. Piston-Hatlen (eds.) *Le français en Amérique du Nord: état présent. Québec*, Québec: Les Presses de l'Université Laval, 261-286.

Nagy, N. & H. Blondeau (2005). 'I think [CP [COMP Ø] [IP that is deleted in Montreal]]', Poster presented at NWAV 34, New York University.

Neumann-Holzschuh, I. & J. Mitko (2018). *Grammaire comparée des français d'Acadie et de Louisiane (GraCoFAL), avec un aperçu sur Terre-Neuve*, Berlin/Boston: Walter de Gruyter.

Nevaci, M. & A. Todi (2009). 'The Grammaticalization of Perfect Auxiliaries in Romanian. Historical and Dialectal Aspects', *Revue roumaine de linguistique* 54: 137-50.

Nyrop, K. (1930). *Grammaire historique de la langue française,* Copenhagen: Gyldendal.

Orlandi, G. (1989). *Il dialetto di Roiate*. Rome: Carucci.

Paillé, M. (2011). 'Les caractéristiques linguistiques de la population du Québec: profil et tendances 1996-2006', *Portrait démolinguistique*, September 2011. Montréal: Office québécois de la langue française, Gouvernement du Québec. Web. http://www.oqlf.gouv.qc.ca/ressources/sociolinguistique/etudes2011/20110909_MPaille.pdf [Accessed April 11, 2017]

Pană Dindelegan, G. (2013) 'The Participle', In G. Pană Dindelegan (ed.) *The Grammar of Romanian*, Oxford: Oxford University Press, 222-232.

Papen, R. (2016). 'Les variétés de français de l'Ouest canadien : un bilan de plus de 60 ans de recherche', Paper presented at *Les français d'ici*, Université de Saint-Boniface, Winnipeg, Manitoba, 7-9 June 2016.

Papen, R. & K. Rottet (1997). 'A Structural sketch of the Cajun French spoken in Lafourche and Terrebonne Parishes', In A. Valdman (ed.) *French and Creole in Louisiana*, New York: Plenum Press, 71-108.

Patch-Neveu, C. (2019). 'Le bilinguisme au Canada demeure un utopie', *Le Devoir* [Montréal], Septembre 18. https://www.ledevoir.com/opinion/lettres/562830/lettre-le-bilinguisme-au-canada-demeure-une-utopie

Penny, R. (2000). *Variation and Change in Spanish*. Cambridge: Cambridge University Press.

Perlmutter, D. (1978). 'Impersonal passives and the unaccusative hypothesis', In *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society* (BLS 4), Berkeley, CA: Berkeley Linguistics Society, 57-89

Perlmutter, D. (1989). 'Multiattachment and the Unaccusative Hypothesis: the perfect auxiliary in Italian'*, Probus* 1: 63-119.

Péronnet, L. (1989). *Le parler acadien du sud-est du Nouveau-Brunswick : éléments grammaticaux et lexicaux*, Berne: Peter Lang.

Picone, M. D. & A. Valdman (2005). 'La situation du français en Louisiane', In A. Valdman, J. Auger & D. Piston-Hatlen (eds.) *Le français en Amérique du Nord*, Québec: Presses de l'Université Laval, 143-168

Pooley, T. (1988). *Grammatical and phonological variation in the working-class French of Roubaix*. Unpublished PhD thesis: University of London.

Poplack, S. (1992). 'The inherent variability of the French subjunctive', In C. Laeufer & T. A. Morgan (eds.) *Theoretical Analyses in Romance Linguistics*, Amsterdam: Benjamins, 235-263.

Poplack, S. (1993). 'Variation theory and language contact', *American Dialect Research*: 251.

Poplack, S. (1997). 'The sociolinguistic dynamics of apparent convergence', In G. Guy, J. Baugh & D. Schiffrin (eds.) *Towards a Social Science of Language*, Amsterdam: Benjamins, 285-309.

Poplack, S. & D. Turpin (1999). 'Does the FUTUR have a future in (Canadian) French?', *Probus* 11.1, 133-164.

Poplack, S. & N. Dion (2004). 'The French future in grammar and speech', Paper presented at *NWAV 33,* Ann Arbor, USA.

Poplack, S. & A. St-Amand (2009). 'Les Récits du français québécois d'autrefois : reflet du parler vernaculaire du 19e siècle', *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 54: 511-546.

Presnukhina, Y. (2016). *Indicateurs de suivi de la situation linguistique au Québec : portrait démolinguistique 1996-2011*. Montréal : Office québécois de la langue française, Gouvernement du Québec. https://www.oqlf.gouv.qc.ca/etudes2016/portrait-demolinguistique-19962011.pdf [Accessed April 11, 2017]

R Core Team (2004-2019). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. URL http://www.R-project.org/

Radio-Canada (2012). 'Le français et l'anglais reculent à Montréal', *Radio-Canada* [Montréal], 24 October. Web. http://ici.radio-canada.ca/nouvelle/584310/recensement-langues-francais [Accessed April 11, 2017]

Rea, B. (2014). *Aspects of Pronoun and Auxiliary Morphology in French, with particular Reference to Spoken Montréal French*. Unpublished MPhil dissertation: University of Oxford.

Rea, B. (2020). 'Chapter 10: A comparative analysis of French auxiliation, with new evidence from Montréal', In S. Wolfe & M. Maiden (eds.) *Variation and Change in Gallo-Romance Grammar*, Oxford: Oxford University Press, 213-240.

Remacle, L. (1956). *Syntaxe du parler wallon de La Gleize. Tome 2: verbes, adverbes, prepositions,* Paris: Les Belles Lettres.

Remysen, W. (2001). *Le sentiment d'insécurité linguistique des Québécois : analyse des attitudes linguistiques dans le discours épilinguistique*. Master's dissertation: Katholieke Universiteit Leuven, 2 vol.

Remysen, W. (2004). 'La variation linguistique et l'insécurité linguistique : le cas du français québécois', In P. Bouchard (ed.) *La variation dans la langue standard. Actes du colloque tenu les 13 et 14 mai 2002 à l'Université Laval dans le cadre du 70e Congrès de l'ACFAS*, Québec: Office québécois de la langue française, 23-36.

Remysen, W. (2018). 'L'insécurité linguistique à l'école : un sujet d'étude et un champ d'intervention pour les sociolinguistes', In N. Vincent & S. Piron (eds.) *La linguistique et le dictionnaire au service de l'enseignement du français au Québec*, Montréal: Nota Bene.

Remysen, W. (2020). 'S'adapter pour se sentir moins insécurisé ? L'(in)sécurité linguistique au Québec sous l'angle de l'accommodation', In J. Lorilleux & V. Feussi (eds.) *L'(in)sécurité linguistique en Francophonies : perspectives interdisciplinaires*, Paris: L'Harmattan (« Espaces discursifs »), 63-75.

Renaud, C. & A.-J. Villeneuve (2008). 'L'alternance des auxiliaires *avoir* et *être* à Chicoutimi-Jonquière', Paper presented at the conference *Les français d'ici* (LFDI 2), Ottawa (Canada), May 23.

Rideout, D. L. (2011). 'Auxiliary Selection in 16th Century French: Imposing Norms in the Face of Language Change', In S. Johanson, E. Hazenberg, and S. Power (eds.) *Proceedings of the 33rd Atlantic Provinces Linguistics Association Memorial University of Newfoundland: Memorial University of Newfoundland Occasional Papers in Linguistics, vol. 2*.

Rideout, D. L. (2013). Avoir *ou* être *dans les formes verbales composées : Conflits, usages et choix des grammairiens dans l'histoire du français de 1500 à 1789*. Ph.D. thesis: Université de Montréal.

Roberge, Y. & N. Rosen (1999). 'Preposition stranding and *que*-deletion in varieties of North American French', *Linguistica Atlantica*, 21: 153-168.

Romaine, S. & Southwest Educational Development Lab., Austin, TX. (1981). *On the Problem of Syntactic Variation A Reply to Beatriz Lavandera and William Labov. Sociolinguistic Working Paper Number 82*. Washington, D.C.

Ronjat, J. (1937). *Grammaire istorique [sic] des parlers provençaux modernes*, Montpellier: Société des langues romanes.

Rosen, C. (1984). 'The interface between semantic roles and initial grammatical relations', In D. Perlmutter & C. Rosen (eds.) *Studies in Relational Grammar* 2, Chicago: Chicago University Press, 38-77

Rousseau, P. & D. Sankoff (1978). *Advances in variable rule methodology. Linguistic variation: models and method*s, ed. by D. Sankoff, New York, NY: Academic Press, 57-69.

Roussel, B. (2016). 'Deux auxiliaires pour le prix d'un: l'alternance entre *avoir* et *être* en français acadien', Paper presented at *Les français d'ici*, Université de Saint-Boniface, Winnipeg, Manitoba, 7-9 June 2016.

Roux, J. (1895). *Grammaire limousine*, Brive: École limousine.

Russo, M. & J. Roberts (1999). 'Linguistic change in endangered dialects: The case of alternation between *avoir* and *être* in Vermont French', *Language Variation and Change* 11, Cambridge: Cambridge University Press, 67-85.

Sammons, O., Nadasdi, T. & R. Mougeon (2015). 'Moving through the past: thirty years of *avoir été* in Ontario French'*, Journal of French Language Studies* 25(3), 397-422.

Sankoff, D. (1975). *VARBRUL version 2*. Unpublished program and documentation.

Sankoff, D. (1988). 'Problems of representativeness', In U. Ammon, N. Dittmar & K. J. Mattheier (eds.) *Sociolinguistics: an International Handbook of the Science of Language and Society,* Berlin: Walter de Gruyter, 899-903.

Sankoff, D., Sankoff, G., Laberge, S. & M. D. Topham (1976). 'Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale', *Cahier de linguistique* 6: 85-125. doi: 10.7202/800043ar.

Sankoff, D. & S. Laberge (1978). 'The linguistic market and the statistical explanation of variability', in D. Sankoff (ed.) *Linguistic Variation: Models and Methods*, New York: Academic Press.

Sankoff, D. & P. Thibault (1981). 'Weak complementarity: Tense and aspect in Montreal French', In B. B. Johns & D. R. Strong (eds.) *Syntactic Change*, collection *Natural Language Studies* 25, Ann Arbor: University of Michigan, 205-216.

Sankoff, D., Tagliamonte, S. & E. Smith (2005). *GoldVarb X: a variable rule application for Macintosh and Windows.*
http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm

Sankoff, G. (1972). 'Above and beyond phonology in variable rules', In C.-J. Bailey & R. W. Shuy (eds.) *New ways of analyzing variation in English*. Washington, D.C.: Georgetown University Press, 44-61.

Sankoff, G. (1974). 'A quantitative paradigm for the study of communicative competence', In R. Bauman & J. Sherzer (eds.) *Explorations in the Ethnography of Speaking,* Cambridge: Cambridge University Press, 18-49.

Sankoff, G. (1980a). 'A quantitative paradigm for the study of communicative competence', In G. Sankoff (ed.) *The Social Life of Language*. Philadelphia: University of Pennsylvania Press, 47-79.

Sankoff, G. (1980b). 'Above and beyond phonology in variable rules', In G. Sankoff (ed.) *The Social Life of Language*. Philadelphia: University of Pennsylvania Press, 81-93.

Sankoff, G. (2004). 'Adolescents, young adults and the critical period: Two case studies from "Seven-Up"', In C. Fought (ed.) *Sociolinguistic Variation*, Oxford: Oxford University Press, 121-139.

Sankoff, G. (2006). 'Age: Apparent Time and Real Time', In K. Brown (chief ed.) *Encyclopedia of Language and Linguistics* II, Oxford: Elsevier, 110-116.

Sankoff, G. (2009). 'Three decades of change and variation in Montréal French', Paper presented at LSRL 39, University of Arizona, 29 March.

Sankoff, G. (2017). 'Before there were corpora: The evolution of the Montreal French project as a longitudinal study', In S. E. Wagner & I. Buchstaller (eds.) *Panel studies of language variation and change,* Abingdon-on-Thames: Routledge, 21-52.

Sankoff, G. (2019). 'Language change across the lifespan: Three trajectory types', *Language* 95(2): 197-229. https://doi.org/10.1353/lan.2019.0029

Sankoff, G. & H. Blondeau (2007). 'Language Change Across the Lifespan: /r/ in Montreal French', *Language* 83(3): 560-588.

Sankoff, G. & H. Cedergren (1971). 'Some results of a sociolinguistic study of Montreal French', In R. Darnell (ed.) *Linguistic Diversity in Canadian Society*, Edmonton: Linguistic Research, Inc., 61-87.

Sankoff, G. & H. Cedergren (1972). 'Sociolinguistic research on French in Montreal', *Lang. Soc. I,* 173-174.

Sankoff, G., R. Sarrasin & H. Cedergren (1971). 'Quelques considérations sur la distribution de la variable *que* dans le français de Montréal', Paper presented at the *Congrès de l'Association canadienne-française pour l'Avancement des Sciences*.

Sankoff, G. & P. Thibault (1977). 'L'alternance entre les auxiliaires *avoir* et *être* en français parlé à Montréal', *Langue française* 34, 81-108.

Sankoff, G. & P. Thibault (1980). 'The alternation between the auxiliaries *avoir* and *être* in Montreal French', In G. Sankoff (ed.) *The Social Life of Language*, Philadelphia: University of Pennsylvania Press, 311-345.

Sankoff, G. & P. Thibault (2011). 'Sur les traces de *m'as* en français québécois de 1971 à 2001', In F. Martineau & T. Nadasdi (eds.) *Le français en contact: hommages à Raymond Mougeon*, Québec: Les Presses de l'Université Laval, 331-354.

Sankoff, G., P. Thibault & S. E. Wagner (2004). 'An apparent time paradox: change in Montréal French auxiliary selection, 1971-1995', Paper presented at NWAV 33, Ann Arbor, MI, 30 September-3 October 2004.

Sankoff, G. & D. Vincent (1977). 'L'emploi productif du *ne* dans le français parlé à Montréal', *Le français moderne* 45: 243-256.

Sankoff, G. & S. E. Wagner (2006). 'Age-grading in retrograde movement: The inflected future in Montréal French', *U. Penn Working Papers in Linguistics,* volume 12.2, 203-216.

Schilling-Estes, N. (2005). 'Language Change in Apparent and Real Time: The Community and the Individual', *University of Pennsylvania Working Papers in Linguistics* 10(2): 219-232.

Schilling, N. (2013). *Sociolinguistic Fieldwork*, Cambridge: Cambridge University Press.

Schumacher, R.A, Pierrehumbert, J. B. & P. LaShell (2014), 'Reconciling Inconsistency in Encoded Morphological Distinctions in an Artificial Language', *Proceedings of the 36th Meeting of the Cognitive Science Society (CogSci2014),* Austin, TX: Cognitive Science Society.

Sicre, P. (1909). *Éléments de grammaire du dialecte de Foix*, Foix: Gadrat aîné.

Smith, J. C. (1989). 'Actualization reanalyzed: evidence from the Romance compound past tenses', In T. J. Walsh (ed.) *Synchronic and Diachronic Approaches to Linguistic Variation and Change* (Georgetown University Round Table on Languages and Linguistics, 1988), Washington, D.C.: Georgetown University Press, 310-325.

Smith, J. C. (1999). 'Markedness and morphosyntactic change revisited: the case of Romance past participle agreement', In S. Embleton, J. E. Joseph & H.-J. Niederehe (eds.) *The Emergence of the Modern Language Sciences: studies in the transition from historical-comparative to structural linguistics in honour of E. F. K. Koerner – Volume 2: methodological perspectives and applications*, Amsterdam & Philadelphia: John Benjamins, 203-216.

Smith, J. C. (2011). 'Change and continuity in form–function relationships', In M. Maiden, J. C. Smith & A. Ledgeway (eds.) *The Cambridge History of the Romance Languages: Structures,* vol. I, Cambridge: Cambridge University Press, 268-317.

Smith, J. C. (2016). 'French and northern Gallo-Romance', In A. Ledgeway & M. Maiden (eds.) *The Oxford Guide to the Romance Languages*, Oxford: Oxford University Press, 292-318.

Sorace, A. (2000). 'Gradients in Auxiliary Selection with Intransitive Verbs', *Language* 76: 859-90.

Sorace, A. (2004). 'Gradience at the Lexicon-Syntax Interface: Evidence from Auxiliary Selection', In A. Alexiadou, E. Anagnostopoulou, & M. Everaert (eds.), *The Unaccusativity Puzzle*, Oxford: Oxford University Press, 243-268.

Statistique Canada (2007). 'RMR de Montréal Carte 2 de 2 : Pourcentage d'anglophones (population ayant l'anglais comme langue maternelle) selon les secteurs de recensement (SR) de 2006', *Recensement du Canada de 2006*. Division de la géographie. Web. http://www12.statcan.gc.ca/census-recensement/2006/as-sa/97-555/maps-cartes/CMA-CT/MontrealLang_Anglophone_fc-2.pdf [Accessed April 11, 2017]

Stéfanini, J. (1962). *La voix pronominale en ancien et moyen français*, Ophrys, Publications de l'Université des Lettres et Sciences Humaines de l'Université de Provence.

Stein, D. (1989). 'Markedness and linguistic change', In O. M. Tomić (ed.) *Markedness in Synchrony and Diachrony*, Berlin: Mouton de Gruyter, 67-86.

Stelling, L.E. (2011). 'The effects of grammatical proscription on morphosyntactic change: Auxiliary variation in Franco-American French', *Arborescences: revue d'études françaises*, n° 1.

Štichauer, P. (2018). 'Lexical splits within periphrasis: mixed perfective auxiliation systems in Italo-Romance', *Morphology* 28(1): 1-23.

Tagliamonte, S. (2006). *Analysing Sociolinguistic Variation (Key Topics in Sociolinguistics)*, Cambridge: Cambridge University Press.

Tailleur, S. (2007). *Avoir et être; une ancienne rivalité? Étude de la sélection de l'auxiliaire de parfait en français classique*. Master's dissertation: Université d'Ottawa.

Thibault, P. (2001). 'Regard rétrospectif sur la sociolinguistique québécoise et canadienne', *Revue québécoise de linguistique* 30 (1): 19-42.

Thibault, P. & G. Sankoff (1997). 'A sociolinguistic study of verb morphology in Montreal French: reanalyzing the alternation between *avoir* and *être*', Paper presented at Pennsylvania Linguistics Colloquium, 22-23 February.

Thibault, P. & D. Vincent (1990). *Un corpus de français parlé: Montréal 84, historique, méthodes et perspectives de recherche,* Québec: Université Laval.

Timberlake, A. (1977). 'Reanalysis and actualization in syntactic change', In C. N. Li (ed) *Mechanisms of Syntactic Change*, Austin & London: University of Texas Press, 141-177.

Tristram, A. (2014). *Variation and Change in French Morphosyntax: The Case of Collective Nouns*, Abindgon: Routledge.

Trudgill, P. (1972). 'Sex, covert prestige and linguistic change in the urban British English of Norwich', *Language in Society* 1(2): 179-195.

Trudgill, P. (1974a). 'The social differentiation of English in Norwich', *Cambridge Studies in Linguistics* 13, Cambridge: Cambridge University Press, X + 211.

Trudgill P. (1974b). 'Linguistic change and diffusion: description and explanation in sociolinguistic dialect geography', *Lang. Soc.* 2, 215-246.
DOI:10.1017/S0047404500004358

Trudgill P. (1986). *Dialects in Contact*, Oxford: Blackwell Publishers.

Tuttle, E. F. (1986). 'The spread of *esse* as universal auxiliary in central Italo-Romance', *Medioevo romanzo* 11, 229-287.

Utheim, U. K. P. (2013). *Auxiliaires et intransitivité en français et dans les langues romanes*. Master's dissertation: Universitetet i Oslo.

Vasseur, G. (1996). *Grammaire des parlers picards du Vimeu (Somme) avec considération spéciale du dialecte de Nibas*, Abbeville: F. Paillart.

Ville de Montréal. (2014). 'Profil socio-démographique : Agglomération de Montréal, juillet 2014', *Montréal en statistiques*, Division de la planification urbaine, Direction de l'urbanisme, Service de la mise en valeur du territoire. Web. http://ville.montreal.qc.ca/pls/portal/docs/PAGE/MTL_STATS_FR/MEDIA/DOCUMENTS/PROFIL_SOCIODEMO_AGGLOMERATION.PDF [Accessed April 12, 2017]

Villeneuve, A. (2016). 'Entre vernaculaire et 'standard' : le français québécois parlé en contexte d'entrevue', Paper presented at *Les français d'ici*, Université de Saint-Boniface, Winnipeg, Manitoba, 7-9 June 2016.

Villers, M. de. (2009). *Multidictionnaire de la langue française,* Montréal: Québec Amérique.

Vincent, D., Laforest, M. & G. Martel (1995). 'Le corpus de Montréal 1995: Adaptation de la méthode d'enquête sociolinguistique pour l'analyse conversationnelle', *Dialangue* 6: 29-46.

Vincent, N. (1982). 'The development of the auxiliaries *habere* and *esse* in Romance', In N. Vincent & M. Harris (eds.) *Studies in the Romance Verb*, London: Croom Helm, 71-97.

Wagner, S. E. (2012). 'Age Grading in Sociolinguistic Theory', *Language and Linguistics Compass* 6(6): 371-382.

Wagner, S. E. & G. Sankoff (2011). 'Age grading in the Montréal French inflected future', *Language Variation and Change* 23(3): 275-313.

Warren, J. (1994). '*Plus ça change, plus c'est pareil*: The case of 'que' in Montreal French', *Culture* 14(2): 39-49.

Weiner J. & W. Labov (1977). 'Constraints on the Agentless Passive', paper presented at the 39th LSA summer meeting, Honolulu.

Weinreich U., Labov W. & Herzog M. (1968). 'Empirical foundations for a theory of language change', In W. P. Lehmann & Y. Malkiel (eds.) *Directions for historical linguistics*, Austin, TX: University of Texas Press, 95-195.

Willis, L. (2000). "*Être ou ne plus être?" Auxiliary alternation in Ottawa-Hull French*. Master's dissertation: Université d'Ottawa.

Wolfe, S. & M. Maiden (eds.) (2020). *Variation and Change in Gallo-Romance Grammar*, Oxford: Oxford University Press.

Zamboni, A. (1998). 'Dal latino tardo al romanzo arcaico: Aspetti diacronico-tipologici della flessione nominale', In P. Ramat & E. Roma (eds.) *Sintassi Storica (Atti del XXX Congresso Internazionale della Società di Linguistica Italiana)*, Rome, 127-46.

Zamboni, A. (2000). *Alle origini dell'italiano: dinamiche e tipologie della transizione dal latino.* Rome: Carocci.

Appendices

9.1. Appendix 3A: Ethical approval

9.1.1. Appendix 3A.1: CUREC 1A approval

SOCIAL SCIENCES & HUMANITIES
INTER-DIVISIONAL RESEARCH ETHICS COMMITTEE

Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD
Tel: +44(0)1865 616576 Fax: +44(0)1865 280467
ethics@socsci.ox.ac.uk

10 February 2016

Ms Beatrice Rea
Department of Linguistics/Philology/Phonetics (LPP) Faculty

Dear Ms Rea,

Research Ethics Approval (CUREC 1A)
Ref No: R44103/RE001

Title: Je m'ai fait mal quand j'ai tombé: Issues regarding auxiliation and reflexivity in spoken Montreal French (1971-2016)

The above application has been considered on behalf of the Social Sciences and Humanities Inter-divisional Research Ethics Committee (IDREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

Should there be any subsequent changes to the project, which raise ethical issues not covered in the original application, you should submit details to the IDREC for consideration.

Yours sincerely,

Claudia Kozeny-Pelling
Research Ethics Manager and Secretary SSH IDREC

cc: Dan Holloway

9.1.2. Appendix 3A.2: Participation information sheet (ENG and FR)

Faculty of Linguistics, Philology & Phonetics
University of Oxford
Clarendon Press Centre
Walton Street
Oxford OX1 2HG

enquiries@ling-phil.ox.ac.uk
Tel: **+44 (0)1865 280400**

Participant Information Sheet

Please take some time to read this information and ask any further questions if anything is unclear. Contact details can be found at the end of this document.

1. **Study title**
   Recounting past events in Montréal French (CUREC 1A R44103/RE001)

2. **Background and aims of the study**
   As part of my research project, I am studying patterns of Montréal French related to the linguistic structure of the way past events are recounted and whether an evolution can be witnessed since the 1970s. The aim of the fieldwork is to record participants speaking Montréal French in the most natural and informal way possible. I am looking for data that reflect the speakers' native judgments of their language. This is not a test, so there are no right or wrong answers!

3. **Who is organising and funding this research?**
   The research for this study is being undertaken by Béatrice Rea who is a DPhil (doctoral) student in the Faculty of Linguistics, Philology & Phonetics at the University of Oxford.

   This project is being supervised by Mr. Martin Maiden, a Professor of the Romance languages, and by Mr. John Charles Smith, a Faculty Lecturer in French Linguistics, at the University of Oxford. This research is funded by the Clarendon Fund as well as by Lady Margaret Hall and has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee (CUREC).

4. **Why have I been invited to take part?**
   This project aims to interview 48 people from a variety of different social backgrounds. You have been invited because you are a native speaker of Montréal French and are between 18 and 55 years old.

5. **Do I have to take part?**
   It is not compulsory to take part in this study and you may ask the researcher questions or ask for more information before deciding whether or not to participate. However, we believe that you have something important to contribute to this study and to further researchers who want to study Montréal French, and we hope that you will contribute. Please note, however, that you are free to withdraw (or withdraw your data) from the study at any time and will not be asked to give any reasons if you do choose to withdraw.

6. **What will happen in the study?**
   If you are happy to take part in the study, you will be asked to sign a consent form and we will arrange an interview with you: it is necessary for all participants to be audio recorded during the interview and you will not be able participate without giving specific consent to this. Ideally, we would like to interview you in your own home but if you would rather hold the interview somewhere more convenient for you, that can also be arranged. The interview will be conducted by Béatrice Rea and will likely last between 1 hour and 1h30 hours per speaker.
   You will first be asked to complete a very short personal information questionnaire (age, gender, occupation, languages spoken, etc.). Then, the interview will consist of a very informal conversation with the researcher and your interview partner around topics related to past events (recent holidays, trips, moving houses, etc.) and attitudes towards Montréal French. In the last part of the interview, you

will be required to read a few sentences and determine whether you feel that they are grammatically correct. This should take less than 5 minutes.

7. What are the possible benefits and potential risks in taking part?

Participating in this study will provide you with an opportunity to relive some of your experiences and memories of your earlier lives. The aim of this project is to create a large database consisting of interviews conducted in Montréal French in order to study linguistics patterns related to the way past events are recounted. These data are very important in allowing researchers to understand which linguistic choices native speakers make and how these choices reflect their identity. This project aims to facilitate this. Moreover, we hope that you will find it rewarding to be able to contribute to public knowledge regarding the way Montréalers speak and to add to a significant collection of archived interviews that will enable researchers to understand more about how the French language works.

We hope that the experience of taking part in this study will be enjoyable for you. However, there is the possibility with work that looks back on earlier life experiences that some upsetting memories may be recalled. You will at no point be obliged to talk about any topic that makes you uncomfortable. If you do feel like re-scheduling or abandoning the interview, do not hesitate to tell the researcher.

8. What happens to the research data provided?

The recorded interviews will be transcribed, stored digitally in conformity with the Data Protection Act, and managed by the researcher. Only the researcher will have access to the interviews and personal information. Publications which come out of this project may use anonymous quotations from your interview. Once the interviews have been deposited at the end of the project, future researchers may also wish to use and quote anonymously your interview in their work. On the consent form you will be asked to confirm that you consent for researchers to use and quote anonymously from your interview. In all works produced using your interview, your name and other personal information will be anonymised and no identifying information will be published.

At the end of this project the audio and digital data collected from interviews with participants will be stored digitally in an encrypted form for future use by the researcher (e.g. post-doctoral work).

9. Will the research be published?

The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives easy access for researchers to the full text of freely available theses, thereby increasing the likely impact and use of that research.

If you agree to participate in this project, the research will be written up as a thesis. All participants are able to request a summary of the research findings should they wish to by contacting the researcher. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The digital online copy of the thesis will be deposited with the Oxford University Research Archive (ORA) and will be published with open access, meaning that it will be available to all internet users.

10. Who do I contact if I have a concern about the study or I wish to complain?

If you have a concern about any aspect of this project, please speak to Béatrice Rea, or to her supervisors Prof. Martin Maiden and Mr. John Charles Smith, who will do her/their best to answer your query. The researchers should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it.

Researcher's contact details: Béatrice Rea
Lady Margaret Hall, Norham Gardens
Oxford, Oxfordshire, United Kingdom OX2 6QA
beatrice.rea@ling-phil.ox.ac.uk +44 (0) 7598 024205

Supervisors' contact details: Professor Martin Maiden
Faculty of Medieval and Modern Languages, Taylor Institution

Oxford, Oxfordshire, United Kingdom OX1 2JF
martin.maiden@mod-langs.ox.ac.uk +44 (0) 1865 270488

Mr. John Charles Smith
St Catherine's College, Manor Road
Oxford, Oxfordshire, United Kingdom OX1 3UJ
johncharles.smith@stcatz.ox.ac.uk +44 (0) 1865 271748

If you remain unhappy or wish to make a formal complaint, please contact the chair of the Research Ethics Committee at the University of Oxford (using the contact details below) who will seek to resolve the matter in a reasonably expeditious manner:

Chair, Social Sciences & Humanities Inter-Divisional Research Ethics Committee
Email: ethics@socsci.ox.ac.uk
Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

Remember, this is not a test, so there are no right or wrong answers!

Thank you for considering participation,

Béatrice Rea
DPhil candidate in Comparative Philology and General Linguistics
Lady Margaret Hall, University of Oxford

Fiche de renseignements pour les participants

Veuillez prendre le temps de lire les informations suivantes et n'hésitez pas à poser des questions.

1. Titre de l'étude

Raconter des événements du passé en français montréalais
(numéro de référence d'approbation éthique de la recherche: CUREC 1A R44103/RE001)

2. Contexte et but de l'étude

Dans le cadre de mon projet de recherche, j'étudie des habitudes de langage reliées à la structure linguistique de la façon dont on raconte des événements du passé en français montréalais, et s'il est possible de noter une évolution depuis les années 1970. Je souhaite collecter des données reflétant les choix spontanés des locuteurs natifs.

Le but de ce travail sur le terrain est donc d'enregistrer des participants parlant le français de Montréal le plus naturellement et le plus familièrement possible. Il ne faut en aucun cas vous forcer pour « bien parler » !

3. Qui organise et finance ce projet de recherche ?

Cette étude est menée par Béatrice Rea, candidate au doctorat à la Faculté de Linguistique, Philologie & Phonétique de l'Université d'Oxford.

Ce projet est dirigé par M. Martin Maiden, professeur de langues romanes, ainsi que par M. John Charles Smith, spécialiste de la linguistique française, à l'Université d'Oxford. Ce projet est financé par le Fonds Clarendon ainsi que par le collège Lady Margaret Hall et a été examiné et approuvé par le Comité d'éthique de recherche de l'Université d'Oxford (CUREC).

4. Pourquoi ai-je été invité(e) à participer ?

Ce projet vise à interviewer 48 locuteurs provenant de divers milieux sociaux. Vous avez été invité(e) parce que le français montréalais est votre langue d'usage et parce que vous avez entre 18 et 65 ans.

5. Suis-je obligé(e) de participer ?

Vous êtes entièrement libre de décider si vous désirez participer et vous pouvez poser autant de questions qu'il est nécessaire ou demander plus d'informations avant de décider de participer. Cependant, nous croyons que votre participation sera une contribution importante à cette étude et pour tous les chercheurs qui s'intéressent au français de Montréal. Veuillez noter que vous pourrez vous retirer du projet (ou retirer vos données) en tout temps sans avoir à fournir d'explication.

6. Qu'arrivera-t-il durant l'entrevue ?

Veuillez noter que l'entrevue devra obligatoirement faire l'objet d'un enregistrement audio. Si vous acceptez de participer, on vous demandera de signer un *formulaire de consentement* avant l'entrevue. Idéalement, nous préfèrerions vous interviewer chez vous, mais si vous préférez un autre endroit tout aussi calme, cela est possible. L'entrevue sera menée par Béatrice Rea et sera d'une durée approximative de 1h-1h30 par locuteur.

On vous demandera d'abord de remplir une courte fiche de renseignements personnels (âge, sexe, occupation, langues parlées, etc.). Ensuite, l'entrevue consistera en une conversation très informelle avec la chercheure et votre partenaire d'entrevue à propos d'événements banals du passé (ce que vous avez fait la journée précédente, vacances récentes, déménagements, etc.). Puis, on vous demandera de remplir un court questionnaire sur vos sentiments à l'égard du français montréalais. Dans la dernière partie de l'entrevue, qui durera moins de 5 minutes, vous devrez déterminer si certaines phrases s'entendent à Montréal. Ce dernier exercice n'a absolument rien à voir avec ce que prescrivent les « règles de grammaire » du français. Je suis uniquement à la recherche de vos intuitions par rapport au français parlé !

7. Quels sont les avantages et les risques potentiels de participer ?

Participer à cette étude vous donnera la chance de vous remémorer et revivre certains événements du passé. Le but de ce projet est de créer une base de données d'entrevues menées en français montréalais afin de pouvoir étudier des habitudes linguistiques propres aux Montréalais et à la façon dont ils racontent des événements passés. Ces données sont cruciales pour permettre aux chercheurs de comprendre les choix linguistiques faits par les locuteurs natifs et comment ces choix reflètent leur identité. Cette étude vise donc à faciliter ce type de recherche. De plus, nous espérons que vous trouverez cette expérience gratifiante et enrichissante puisqu'elle vous permettra d'élargir les

connaissances sur la façon dont parlent les Montréalais et de contribuer à la collection d'entrevues archivées qui permettront aux chercheurs d'en apprendre plus sur les subtilités de la langue française.

Nous espérons que vous prendrez plaisir à vous faire interviewer, mais à aucun moment n'êtes-vous obligé(e) de discuter de sujets qui vous mettent mal à l'aise. Si vous désirez reporter ou abandonner l'entrevue à tout moment, n'hésitez pas à le dire à la chercheure.

8. Qu'arrivera-t-il aux données recueillies ?

Les entrevues enregistrées seront ensuite transcrites, stockées numériquement conformément à la Loi sur la protection des données et gérées par la chercheure. Les directeurs de thèse et la chercheure seront les seules personnes qui auront accès aux données linguistiques et aux renseignements personnels pendant la durée de ce projet. Une fois que le projet sera complété, les entrevues pourront être partagées avec d'autres linguistes, sur demande, qui à leur tour pourront utiliser des citations anonymes de votre entrevue. Sur le *formulaire de consentement*, vous devrez confirmer que vous consentez à ce que les chercheurs puissent utiliser et citer de façon anonyme votre entrevue. Il est possible que les publications tirées de ce projet utilisent des citations anonymes de votre entrevue et dans tous les travaux qui utiliseront votre entrevue, les noms seront changés, aucun renseignement personnel et aucunes informations permettant de vous identifier ne seront publiés.

À la fin de ce projet, les données audio et personnelles des entrevues seront stockées de façon numérique et cryptée en vue d'une utilisation ultérieure par la chercheure (ex. projet postdoctoral).

9. Le projet de recherche sera-t-il publié ?

L'Université d'Oxford s'engage à disséminer sa recherche au bénéfice de la société et de l'économie et, afin d'honorer cet engagement, elle a créé des archives en ligne pour les documents de recherche. Ces archives contiennent entre autres des copies numériques des thèses de doctorat soutenues avec succès dans le cadre d'un programme d'études supérieures de l'Université d'Oxford. Le dépôt de ces archives en ligne donne aux chercheurs l'accès complet aux thèses disponibles, ce qui maximise l'impact potentiel de ces recherches.

Si vous acceptez de participer à ce projet, la recherche sera rédigée sous forme de thèse. Tous les participants peuvent demander un résumé des résultats de recherche s'ils le souhaitent, en contactant la chercheure. Si la thèse est soutenue avec succès, elle sera déposée en version imprimée et en ligne dans les archives de l'Université d'Oxford, pour faciliter son utilisation dans des travaux de recherche à venir. La thèse sera publiée en ligne auprès de la Oxford University Research Archive (ORA) en accès libre, ce qui signifie qu'elle sera accessible à tous les internautes.

10. Qui dois-je contacter si j'ai des inquiétudes au sujet de cette étude ou si je veux porter plainte ?

Si vous désirez faire part d'une préoccupation à propos de ce projet, veuillez directement contacter Béatrice Rea, ou ses directeurs de thèse M. Martin Maiden et M. John Charles Smith, qui feront de leur mieux pour répondre à vos questions. Les chercheurs traiteront votre demande dans les plus brefs délais, dans une limite de 10 jours ouvrables, et vous indiqueront comment ils comptent remédier au problème.

Coordonnées de la chercheure: Béatrice Rea
Lady Margaret Hall, Norham Gardens
Oxford, Oxfordshire, United Kingdom OX2 6QA
beatrice.rea@ling-phil.ox.ac.uk +44 (0) 7598 024205

Coordonnées des directeurs de thèse: Professor Martin Maiden
Trinity College, Broad Street
Oxford, Oxfordshire, United Kingdom OX1 3BH
martin.maiden@mod-langs.ox.ac.uk +44 (0) 1865 270488

Mr. John Charles Smith
St Catherine's College, Manor Road
Oxford, Oxfordshire, United Kingdom OX1 3UJ
johncharles.smith@stcatz.ox.ac.uk +44 (0) 1865 271748

Si vous demeurez insatisfait(e) ou souhaitez déposer une plainte officielle, veuillez contacter le président du Comité d'éthique de recherche de l'Université d'Oxford qui tentera de résoudre le problème le plus rapidement possible :

Chair, Social Sciences & Humanities Inter-Divisional Research Ethics Committee
Courriel: ethics@socsci.ox.ac.uk
Adresse: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

L'objectif de ce travail sur le terrain est d'enregistrer des participants parlant le français de Montréal avec le plus de naturel possible. Gardez à l'esprit que ceci n'est en aucun cas un test – il n'y a donc ni bonnes ni mauvaises réponses !

Merci d'envisager de participer à ce projet,

Béatrice Rea
Candidate au doctorat en philologie comparée et linguistique générale
Lady Margaret Hall, University of Oxford

9.1.3. Appendix 3A.3: Consent form templates (ENG and FR)

Participant Consent Form

STUDY TITLE: *Recounting past events in Montréal French*

RESEARCHER DETAILS: Béatrice Rea (doctoral student), Lady Margaret Hall, University of Oxford
beatrice.rea@ling-phil.ox.ac.uk +44 (0) 7598 024205

PURPOSE OF STUDY: As part of my research project, I am studying patterns of Montréal French related to the linguistic structure of the way past events are recounted and whether an evolution can be witnessed since the 1970s. The aim of the fieldwork is to record participants speaking Montréal French in the most natural and informal way possible. I am looking for data that reflect the speakers' native judgments of their language. This is not a test, so there are no right or wrong answers!

Declaration: I, the interviewee, confirm that	Participant initials
1. I have read the participant information sheet, have had the opportunity to ask questions and have received satisfactory answers.	
2. I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.	
3. I understand that my participation is voluntary and that I am free to withdraw myself or my data at any time, without giving any reason, and without any adverse consequences.	
4. I understand who will have access to the personal data provided.	
5. I understand how personal data will be stored (according to the Data Protection Act); and what will happen to the data at the end of the project.	
6. I understand how the research will be written up and published.	
7. I understand how to raise concerns or make a complaint.	
8. I consent to being audio recorded.	
9. I understand that audio recordings and anonymous quotes from the interview might be used in research outputs, ex. doctoral thesis, academic conferences, journal publications, research archives.	
10. I agree to take part in the study.	

Name of participant: _____

Signature: _____**date:**_____

Name of researcher: _____

Signature:_____**date:**_____

Formulaire de consentement

TITRE DE L'ÉTUDE : *Raconter des événements du passé en français montréalais*

COORDONNÉES DE LA CHERCHEURE : Béatrice Rea (candidate au doctorat)
Lady Margaret Hall, University of Oxford
beatrice.rea@ling-phil.ox.ac.uk +44 (0) 7598 024205

BUT DE L'ÉTUDE : J'étudie des habitudes de langage en français montréalais reliées à la structure linguistique de la façon dont on raconte des événements du passé, et s'il est possible de noter une évolution depuis les années 1970. Le but de ce travail sur le terrain est d'enregistrer des participants parlant le français de Montréal le plus naturellement et le plus familièrement possible, car je souhaite collecter des données qui reflètent les choix spontanés des locuteurs natifs.
Ceci n'est pas un test, il n'y a donc ni bonne ni mauvaise réponse !

Déclaration : Je, l'interviewé(e), confirme que	Initiales de l'interviewé(e)
11. J'ai lu la fiche de renseignements pour les participants et, si nécessaire, j'ai eu l'occasion de poser des questions et ai reçu des réponses satisfaisantes.	
12. Je comprends que ce projet a fait l'objet d'un examen par le Comité d'éthique de recherche de l'Université d'Oxford et qu'il a été approuvé sur le plan éthique.	
13. Je comprends que ma participation est volontaire et que je suis libre de me retirer ou de retirer mes données du projet en tout temps, sans explication, et sans conséquences néfastes.	
14. Je comprends qui aura accès aux renseignements personnels que j'aurai fournis.	
15. Je comprends comment mes renseignements personnels seront stockés (en vertu de la Loi sur la protection des données) et ce qu'il arrivera aux données à la fin du projet.	
16. Je comprends comment la recherche sera rédigée et publiée.	
17. Je comprends comment soulever des préoccupations ou porter plainte.	
18. Je consens à ce que l'entrevue fasse l'objet d'un enregistrement audio.	
19. Je comprends que cet enregistrement sera utilisé à des fins de recherche et que des citations anonymes de l'entrevue pourront potentiellement être utilisées dans la thèse de doctorat, des colloques universitaires, des publications dans des revues savantes et des archives de recherche.	
20. J'accepte de participer au projet de recherche.	

Nom de l'interviewé(e) :_____

Signature :_____ **date :**_____

Nom de la chercheure :_____

Signature :_____ **date :**_____

9.2. Appendix 3B: Personal information questionnaire (ENG and FR)

Personal information questionnaire

1. Name: _____

2. Age: _____ **3. Gender**: M / F (please circle)

4. Place of birth or childhood: _____

5. Years spent in the Greater Montréal area: _____

6. Occupation: _____

7. Last school diploma obtained (number of years spent in school):

_____

8. If you have a life partner, what is his or her occupation?

_____

9. What was the main occupation of your parents when you were a child?

Father: _____ Mother:_____

10. Do you speak English at work or at home on a daily basis?

Yes / No (please circle)

11. Other language(s) spoken fluently and regularly:

_____

12. Have you had contact, over an extended period of time, with speakers of the following varieties of French: from France, Belgium, Switzerland, Haiti, Senegal, Benin, etc.?

Yes / No (please circle)

If so, which varieties were they? _____

13. In which neighbourhood of the Greater Montréal area have you lived the longest and in which one do you currently live?

_____

Renseignements personnels

1. Nom : _____

2. Âge : _____ **3. Sexe** : H / F

4. Ville de naissance (ou de votre enfance) : _____

5. Nombre d'années passées dans le Grand Montréal : _____ **ans**

6. Profession/Occupation principale : _____

7. Dernier diplôme obtenu
(ou nombre d'années de scolarisation) : _____

8. Si vous avez un(e) conjoint(e), quelle est sa profession ou son occupation principale?

_____

9. Quelle était la profession ou l'occupation principale de vos parents quand vous étiez enfant ?

Père : _____ Mère : _____

10. Utilisez-vous l'anglais quotidiennement au travail ou à la maison depuis longtemps ?
Oui / Non

11. Autre(s) langue(s) parlée(s) très couramment :

_____

12. Avez-vous déjà côtoyé, sur une période prolongée, des locuteurs des variétés de français suivantes : de la France, de la Belgique, de la Suisse, d'Haïti, du Sénégal, du Bénin, etc.?

Oui / Non

Si oui, de quelle(s) variété(s) s'agit-il ? _____

13. Dans quelle municipalité/quel quartier du Grand Montréal avez-vous habité le plus longtemps, et dans quelle municipalité/quel quartier habitez-vous maintenant?

9.3. Appendix 3C: Grammaticality judgement task

Acceptability of *avoir* selection with intransitive *Ê*-verbs and pronominal verbs

Green sentences: Sentences without the variable under study
Red sentences: 'Control' sentences with the *être* variant
Black sentences: Sentences with the *avoir* variant

1. *On a rangé la maison pour les invités.*
   'We cleaned the house for the guests'

2. *Une chance que la bourse a remonté !*
   'Fortunately the stock market went back up again!'

3. *Il est survenu à l'improviste.*
   'He appeared unexpectedly'

4. *J'ai allé au cinéma hier.*
   'I went to the cinema yesterday'

5. *Les vacances sont déjà finies… Ça a tellement passé vite !*
   'The holiday are already over… It went by so quickly!'

6. *Tu as porté ta robe préférée au mariage de Julie.*
   'You wore your favorite dress to Julie's wedding'

7. *On est entrés dans la pièce sans faire de bruit.*
   'We entered the room without making noise'

8. *J'ai né à Montréal en 1982.*
   'I was born in Montréal in 1982'

9. *Il m'a prêté sa voiture pour la fin de semaine.*
   'He lent me his car for the weekend'

10. *J'ai sorti avec elle pendant deux ans à peu près.*
    'I dated her for about two years'

11. *On a resté neuf jours à Paris, puis après on a pris le train pour Londres.*
    'We stayed nine days in Paris, and then after that we took the train to London'

12. *C'est vous qui êtes venus ce matin ?*
    'Are you the one(s) who came this morning?'

13. *On a retourné en arrière parce qu'on avait oublié le chien à la maison.*
    'We turned back because we had forgotten the dog at home'

14. *Tu as démarré l'auto avant de l'avoir complètement déneigée.*
    'You started the car before having completely cleared the snow around it'

15. *Je m'avais déguisé en Père Noël pour les enfants.*
'I had dressed up as Santa Claus for the children'

16. *Elle est morte sur le coup.*
'She died instantly'

17. *Il a parvenu à le convaincre d'arrêter de fumer.*
'He managed to convince him to quit smoking'

18. *On est partis avant la fin du spectacle.*
'We left before the end of the show'

19. *Tu peux pas m'embrasser sans t'avoir brossé les dents !*
'You can't kiss me before having brushed your teeth!'

20. *Ils ont transporté toutes les boîtes d'une maison à l'autre en quelques heures à peine.*
'They carried all the boxes from one house to the other in a few hours only'

21. *Il a devenu tellement fâché quand il a appris ce qui s'est passé.*
'He became so angry when he found out what happened'

22. *J'ai changé la couche du bébé.*
'I changed the baby's diaper'

23. *Mon grand-père a mouru le mois passé.*
'My grand-father died last month'

24. *On est devenus paresseux.*
'We became lazy'

25. *Tu as rapporté son cadeau au magasin parce qu'elle avait reçu la même chose pour Noël.*
'You returned her gift to the store because she had received the same thing for Christmas'

26. *Il a beaucoup changé depuis son divorce.*
'She has changed since her divorce'

27. *J'ai venu aussitôt que j'ai appris la mauvaise nouvelle.*
'I came as soon as I found out the bad news'

28. *Tu as marché jusqu'au métro.*
'You walked all the way to the subway'

29. *J'ai comme eu l'impression d'avoir retombé en enfance quand j'ai vu notre ancienne maison !*
'I sort of had the feeling that I had been brought back in my childhood when I saw our old house!'

30. *C'était trop dangereux, donc j'ai sorti de là le plus vite que j'ai pu.*
'It was too dangerous, so I got out of there as quickly as I could'

31. *On est allés au parc hier.*
'We went to the park yesterday'

32. *Elles ont terminé la course, même si elles étaient très fatiguées.*
'They finished the race, even though they were very tired'

33. *Elle a des gros problèmes de genou, je crois pas qu'elle aurait descendu au sous-sol sans sa canne.*
'She has important knee problems, I don't think that she would have gone down to the basement without her cane'

34. *Tu es revenue trop tard.*
'You came back too late'

35. *Je m'ai mis à l'aise quand j'ai vu que personne s'était habillé chic.*
'I started to get comfortable when I saw that no one had dressed up'

36. *Ils ont demeuré à Outremont pendant 15 ans.*
'They resided in Outremont for 15 years'

37. *Tu as mis la lasagne du four.*
'You put your lasagna in the oven'

38. *On a arrivé presque en même temps qu'eux.*
'We arrived almost at the same time as them'

39. *– C'est de ta faute l'accident d'auto ?*
– Ben non, c'est lui qui m'a rentré dedans !
'– Is the car accident your fault?
– Of course not, he is the one who hit me!'

40. *T'es resté chez elle pour la réconforter.*
'You stayed at her place to comfort her'

41. *Elles ont développé une belle complicité.*
'They developed a beautiful complicity'

42. *J'ai monté jusqu'en haut de la montagne. C'était toute une randonnée !*
'I climbed all the way to the top of the mountain. It was quite a hike!'

43. *Il a oublié d'aiguiser ses patins.*
'He forgot to sharpen his skates'

44. *J'ai rentré par la fenêtre parce que j'avais oublié mes clés.*
'I got in through the window because I had forgotten my keys'

45. *Tu penses-tu qu'elle a aimé son cadeau ?*
'Do you think that she liked her gift?'

46. *Durant la tempête de verglas, l'arbre a tombé sur les fils électriques.*
'During the ice storm, the tree fell on the electric wires'

47. *Ils ont assemblé tous ses meubles en moins de trois heures.*
'They assembled all her furniture in less than three hours'

48. *J'avais vraiment mal parce que je m'avais foulé la cheville en tombant.*
'I was in a lot of pain because I had sprained my ankle during a fall'

49. *Est-ce que vous êtes née pendant la guerre ?*
'Were you born during the war?'

50. *On a déménagé à Sherbrooke y'a deux ans.*
'We moved to Sherbrooke two years ago'

51. *L'accident a survenu sans qu'on puisse faire quoi que ce soit.*
'The accident happened without us being able to do anything'

52. *On a recouvert les meubles de jardin quand il a commencé à pleuvoir.*
'We covered the garden furniture when it started to rain'

53. *T'es retourné le voir une dernière fois.*
'You went back to see him one last time'

54. *Elle a revenu chez ses parents après ses études.*
'She moved back with her parents after her studies'

55. *J'ai parti de là dès que je l'ai vu arriver.*
'I left that place as soon as I saw him arrive'

56. *J'ai tapé la mauvaise adresse sur mon GPS.*
'I put in the wrong adress in my GPS'

57. *Après avoir pris ma douche, je m'ai habillé en vitesse.*
'After having showered, I quickly got dressed'

58. *Ils sont arrivés juste à temps.*
'They arrived just in time'

59. *Elles ont intervenu deux fois pendant le débat.*
'They intervened twice during the debate'

60. *Tu as commencé à étudier à la dernière minute.*
'You started to revise at the last minute'

9.4. Appendix 3D: Linguistic attitudes questionnaire (followed by an English translation)

Nom : _____

Questionnaire sur les attitudes linguistiques

1. Dans la vie de tous les jours, vous identifiez-vous comme … ?

A) Canadien(ne)

B) Québécois(e)

C) Montréalais(e)

D) Aucune de ces réponses. Je m'identifie plutôt comme

_____

2. Selon moi, le français que je parle diffère du français standard au niveau _____. (Encerclez toutes les réponses qui vous semblent appropriées)

A) De la prononciation (ou de l'accent)

B) Du vocabulaire (ou des expressions et des mots utilisés)

C) De la langue parlée

D) De la langue écrite

E) Aucune de ces réponses, il n'y a pas de différences significatives.

3. Selon moi, le français que la majorité des Québécois parlent à la maison diffère du français standard au niveau _____. (Encerclez toutes les réponses qui vous semblent appropriées)

A) De la prononciation (ou de l'accent)

B) Du vocabulaire (ou des expressions et des mots utilisés)

C) De la langue parlée

D) De la langue écrite

E) Aucune de ces réponses, il n'y a pas de différences significatives.

I. Êtes-vous d'accord avec les affirmations suivantes ?

4. Il est important de continuer à parler français au Québec.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

5. La Loi 101, qui fait du français « la seule langue de communication officielle de l'État québécois et des entreprises qui font des affaires au Québec », doit continuer à être appliquée.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

6. Je parle mieux français que la moyenne des Québécois.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord : je parle français comme la moyenne des Québécois.

D) Je suis tout à fait en désaccord : je parle moins bien français que la moyenne des Québécois.

7. Afin de protéger la culture québécoise, il est important de parler le français québécois.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

8. Le français qu'on retrouve, par exemple, au *Téléjournal* de Radio-Canada et/ou celui qui est prescrit par les ouvrages de grammaire a plus de valeur pour moi que le français québécois.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord : ils ont autant de valeur.

D) Je suis tout à fait en désaccord : le français québécois a plus de valeur pour moi.

E) Il n'y a aucune différence entre les deux.

9. C'est une fierté de parler le français québécois.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

10. Le français québécois est plus expressif que le français standard.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord : ils sont autant expressifs.

D) Je suis tout à fait en désaccord : le français standard est plus expressif.

E) Il n'y a aucune différence entre les deux.

11. Bien maîtriser le français est important.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

12. Le français standard est plus acceptable grammaticalement que le français québécois.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

E) Il n'y aucune différence entre les deux.

13. Plus on est instruit, moins on utilise le français québécois.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

14. Je suis fier(ère) de faire partie de la société québécoise.

A) Je suis tout à fait d'accord.

B) Je suis plutôt d'accord.

C) Je suis plutôt en désaccord.

D) Je suis tout à fait en désaccord.

15. Les francophones parlent mieux français à Montréal que dans le reste du Québec.

    A) Je suis tout à fait d'accord.

    B) Je suis plutôt d'accord.

    C) Je suis plutôt en désaccord : ils parlent français aussi bien que dans le reste du Québec.

    D) Je suis tout à fait en désaccord : ils parlent français moins bien que dans le reste du Québec.

16. Lorsqu'on interagit avec des Français (citoyens de la France) établis au Québec, il est important de faire des efforts pour mieux parler français afin qu'ils nous comprennent.

    A) Je suis tout à fait d'accord.

    B) Je suis plutôt d'accord.

    C) Je suis plutôt en désaccord.

    D) Je suis tout à fait en désaccord.

17. Lorsqu'on interagit avec des anglophones du Québec, il est important de communiquer en français.

    A) Je suis tout à fait d'accord.

    B) Je suis plutôt d'accord.

    C) Je suis plutôt en désaccord.

    D) Je suis tout à fait en désaccord.

18. Lorsqu'on voyage en France, il est important de faire des efforts pour mieux parler français pour que les Français nous comprennent.

    A) Je suis tout à fait d'accord.

    B) Je suis plutôt d'accord.

    C) Je suis plutôt en désaccord.

    D) Je suis tout à fait en désaccord.

II. Veuillez encercler la réponse qui correspond le mieux à votre opinion.

19. Que pensez-vous de l'état du français tel qu'il est parlé dans votre quartier/municipalité ?

    A) Il est de très bonne qualité.

    B) Il est de bonne qualité.

    C) Il est de mauvaise qualité.

    D) Il est de très mauvaise qualité.

20. Que pensez-vous de l'état du français tel qu'il est parlé dans votre milieu de travail, ou votre milieu d'études si vous êtes toujours étudiant(e) ?

    A) Il est de très bonne qualité.

    B) Il est de bonne qualité.

    C) Il est de mauvaise qualité.

    D) Il est de très mauvaise qualité.

21. Que pensez-vous de l'état du français tel qu'il est parlé à Montréal en général ?

    A) Il est de très bonne qualité.

    B) Il est de bonne qualité.

    C) Il est de mauvaise qualité.

    D) Il est de très mauvaise qualité.

22. Que pensez-vous de l'enseignement du français dans les écoles primaires et secondaires du Québec ?

    A) Il est de très bonne qualité.

    B) Il est de bonne qualité.

    C) Il est de mauvaise qualité.

    D) Il est de très mauvaise qualité.

23. Que pensez-vous du français utilisé dans les médias québécois (télévision, radio, journaux, etc.) ?

    A) Il est de très bonne qualité.

    B) Il est de bonne qualité.

    C) Il est de mauvaise qualité.

    D) Il est de très mauvaise qualité.

24. À votre avis, la musique québécoise est en général _____.

    A) De très bonne qualité.

    B) De bonne qualité.

    C) Elle est de mauvaise qualité.

    D) Elle est de très mauvaise qualité.

25. À votre avis, les films québécois sont en général _____.

A) De très bonne qualité.

B) De bonne qualité.

C) Ils sont de mauvaise qualité.

D) Ils sont de très mauvaise qualité.

26. À votre avis, la littérature québécoise est en général _____.

A) De très bonne qualité.

B) De bonne qualité.

C) Elle est de mauvaise qualité.

D) Elle est de très mauvaise qualité.

27. À votre avis, qu'est-ce qui pourrait être amélioré dans le français parlé à Montréal ? Cochez toutes les réponses qui vous semblent appropriées.

A) La prononciation

B) Le vocabulaire

C) La structure des phrases

D) L'utilisation d'anglicismes

E) Le niveau de politesse

F) Rien du tout

Name : _____

Linguistic attitudes questionnaire

1. In your everyday life, do you identify as a… ?

A) Canadian

B) Quebecer

C) Montréaler

D) None of these answers. I identify as a _____

2. In my opinion, the French that I speak differs from standard French in terms of _____. (Circle all the answers that apply)

A) The pronunciation (or the accent)

B) The vocabulary (or the expressions and the words used)

C) The spoken language

D) The written language

E) None of these answers, because there are no significant differences.

3. In my opinion, the French that the majority of Quebecers speak at home differs from standard French in terms of _____. (Circle all the answers that apply)

A) The pronunciation (or the accent)

B) The vocabulary (or the expressions and the words used)

C) The spoken language

D) The written language

E) None of these answers, because there are no significant differences.

I. Do you agree with the following statements?

4. It is important to keep on speaking French in Québec.

A) I strongly agree.

B) I tend to agree.

C) I tend to disagree.

D) I strongly disagree.

5. Bill 101, which ensures that French is the sole language of the state of Québec and of the Québec businesses, should continue to be applied.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree.

    D) I strongly disagree.

6. I speak French better than the average Quebecer.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree: I speak French like the average Quebecer.

    D) I strongly disagree: I speak French worse than the average Quebecer.

7. In order to protect Québécois culture, it is important to speak Québec French.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree.

    D) I strongly disagree.

8. In my opinion, the French that is spoken, for example, on the *Téléjournal* of Radio-Canada and/or the one that is prescribed in grammar books has more value than Québec French.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree: the two have the same value.

    D) I strongly disagree: Québec French has more value.

    E) There is no difference between the two.

9. I am proud to speak Québec French.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree.

    D) I strongly disagree.

10. Québec French is more expressive than standard French.

A) I strongly agree.

B) I tend to agree.

C) I tend to disagree: they are equally expressive.

D) I strongly disagree: standard French is more expressive.

E) There is no difference between the two.

11. Having an excellent knowledge of French is important.

A) I strongly agree.

B) I tend to agree.

C) I tend to disagree.

D) I strongly disagree.

12. Standard French is more grammatically acceptable than Québec French.

A) I strongly agree.

B) I tend to agree.

C) I tend to disagree.

D) I strongly disagree.

E) There is no difference between the two.

13. The more educated one is, the less one tends to use Québec French.

A) I strongly agree.

B) I tend to agree.

C) I tend to disagree.

D) I strongly disagree.

14. I am proud to belong to Québec society.

A) I strongly agree.

B) I tend to agree.

C) I tend to disagree.

D) I strongly disagree.

15. Francophones speak better French in Montréal than in the rest of Québec.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree: they speak French just as well as in the rest of Québec.

    D) I strongly disagree: they speak French worse than in the rest of Québec.

16. When one interacts with French people (citizens of France) that live in Québec, it is important to make efforts to speak French better so that they can understand us.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree.

    D) I strongly disagree.

17. When one interacts with Québec Anglophones, it is important to communicate in French.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree.

    D) I strongly disagree.

18. When one travels to France, it is important to make efforts to speak French better so that French people can understand us.

    A) I strongly agree.

    B) I tend to agree.

    C) I tend to disagree.

    D) I strongly disagree.

II. Please circle the answer that represents your opinion as closely as possible.

19. What is your opinion of the state of French as it is spoken in the neighbourhood/ municipality where you live?

    A) It is of excellent quality.

    B) It is of good quality.

    C) It is of poor quality.

    D) It is of very poor quality.

20. What is your opinion of the state of French as it is spoken in your workplace or at your school?

A) It is of excellent quality.

B) It is of good quality.

C) It is of poor quality.

D) It is of very poor quality.

21. What is your opinion of the state of French as it is spoken in Montréal generally?

A) It is of excellent quality.

B) It is of good quality.

C) It is of poor quality.

D) It is of very poor quality.

22. What is your opinion of French instruction in elementary and secondary schools in Québec?

A) It is of excellent quality.

B) It is of good quality.

C) It is of poor quality.

D) It is of very poor quality.

23. What is your opinion of the French used in Québécois media (television, radio, newspapers, etc.)?

A) It is of excellent quality.

B) It is of good quality.

C) It is of poor quality.

D) It is of very poor quality.

24. In your opinion, Québécois music is generally _____.

A) Of excellent quality.

B) Of good quality.

C) Of poor quality.

D) Of very poor quality.

25. In your opinion, Québécois films are generally _____.

A) Of excellent quality.

B) Of good quality.

C) Of poor quality.

D) Of very poor quality.

26. In your opinion, Québécois literature is generally _____.

A) Of excellent quality.

B) Of good quality.

C) Of poor quality.

D) Of very poor quality.

27. In your opinion, which elements of spoken Montréal French could be improved? (Circle all the answers that apply)

A) Its pronunciation

B) Its vocabulary

C) The structure of its sentences

D) The use of Anglicisms

E) The level of politeness

F) Nothing at all.

9.5. Appendix 4A: Model 1 for intransitive verbs with three distinct socioprofessional statuses (no convergence)

MODEL 1	**Multivariate analysis of the likelihood of *avoir* selection (vs *être*) with 3 SPS levels (no convergence)**		
Model formula: Variant.collapsed ~ Animacy + Hodiernal.recoded +Intervening.element + Pronoun.before.aux + SPS + Tense + Transitive.use + (1 | Code.name) + (1 | Verb.collapsed)			
Input probability	0.273		
Total rate	10,8%		
Total N	2309		
AIC	897.142		
R^2	0.682		
Deviance	873.142		
Significant factors	**Factor weights**	**% *avoir***	**Total N**
Animacy* **	*p* = 4.51e-07		
Inanimate	0.671	23,4	214
Animate	0.329	9,5	2095
Pronoun before auxiliary**	*p* = 1.79e-03		
Pronoun before aux	0.717	20	55
No pronoun before aux	0.283	10,6	2254
Hodiernal action**	*p* = 6.86e-03		
Pre-hodiernal	0.614	12,3	1909
Hodiernal	0.386	3,5	400
Intervening element (between aux and past participle)**	*p* = 8.63e-03		
Intervening element	0.597	18,1	182
No intervening element	0.403	10,2	2127
Tense*	*p* = 0.0115		
Conditionnel passé	0.749	28	25
Passé composé	0.478	10,9	2109
Plus-que-parfait	0.268	7,4	175
Transitive use*	*p* = 0.0118		
Transitive use allowed	0.826	0,25	734
Transitive use not allowed	0.174	4,3	1575
Socioprofessional status*	*p* = 0.0129		
Low	0.653	14,9	698

High	**0.441**	8,6	770
Mid	**0.402**	9,4	841
Speaker (48)	Random		
Lexical item (13)	Random		

Table 9.1 *Shiny Rbrul* (Johnson 2017) results for Model 1: statistical effects of the significant factor groups on auxiliary alternation in the 2016 Montréal data, with three distinct socioprofessional statuses, by *p* value and factor weights (no convergence)

9.6. Appendix 4B: *Avoir* tokens by speaker and by lexical item

Categorical uses with *avoir* are highlighted in bold.

Speaker pseudonym by decreasing order of *avoir*-selection rate	Verbs conjugated with *avoir* by decreasing order of overall *avoir*-selection rate	N of *avoir*/ Total N	% *avoir*
Philippe	***passer***	**1 / 1**	**100**
	tomber	**2 / 2**	**100**
	rentrer	**1 / 1**	**100**
	rester	**1 / 1**	**100**
	retourner	**1 / 1**	**100**
Linda	***déménager***	**1 / 1**	**100**
	tomber	1 / 2	50
	descendre	1 / 2	50
	rentrer	1 / 2	50
	sortir	1 / 3	33,3
	rester	**3 / 3**	**100**
	retourner	**1 / 1**	**100**
Yves	***déménager***	**1 / 1**	**100**
	tomber	**1 / 1**	**100**
	descendre	3 / 5	60
	rentrer	1 / 2	50
	sortir	1 / 2	50
	partir	1 / 7	14,3
Sylvain	***déménager***	**3 / 3**	**100**
	passer	**2 / 2**	**100**
	tomber	1 / 3	33,3
	rentrer	**1 / 1**	**100**
	sortir	2 / 3	66,7
	rester	1 / 2	50
	retourner	**1 / 1**	**100**
	arriver	1 / 16	6,3
Maxime	***déménager***	**1 / 1**	**100**
	passer	**1 / 1**	**100**
	rentrer	**2 / 2**	**100**
	partir	1 / 3	33,3

Sara	***déménager***	**1 / 1**	**100**
	passer	**2 / 2**	**100**
	monter	**2 / 2**	**100**
	tomber	1 / 3	33,3
	descendre	**1 / 1**	**100**
	rentrer	1 / 9	11,1
	sortir	2 / 3	66,7
Joël	***déménager***	**3 / 3**	**100**
	passer	**1 / 1**	**100**
	tomber	**1 / 1**	**100**
	rentrer	**1 / 1**	**100**
	retourner	**1 / 1**	**100**
Sophie	*déménager*	5 / 6	83,3
	passer	3 / 6	50
	monter	**4 / 4**	**100**
	tomber	2 / 3	66,7
	rentrer	1 / 8	12,5
	aller	1 / 25	4
Gaëtan	***passer***	**7 / 7**	**100**
	tomber	4 / 8	50
	retourner	1 / 2	50
Jacynthe	***descendre***	**1 / 1**	**100**
	tomber	1 / 2	50
	sortir	**1 / 1**	**100**
	retourner	1 / 2	50
Rachel	***déménager***	**1 / 1**	**100**
	passer	**1 / 1**	**100**
	monter	2 / 3	66,7
	tomber	1 / 4	25
	rentrer	2 / 6	33,3
	sortir	1 / 3	33,3
	rester	1 / 5	20
	partir	1 / 8	12,5
Dominic	*déménager*	4 / 5	80
	passer	2 / 3	66,7
	monter	1 / 1	100
	tomber	**3 / 3**	**100**
	rester	1 / 3	33,3

Steve	*déménager*	2 / 3	66,7
	rentrer	1 / 2	50
	rester	**3 / 3**	**100**
	arriver	1 / 9	11,1
Jean-François	***déménager***	**1 / 1**	**100**
	monter	**1 / 1**	**100**
	rentrer	4 / 6	66,7
Alexis	***passer***	**2 / 2**	**100**
	tomber	2 / 5	40
	descendre	**1 / 1**	**100**
Julie	***déménager***	**6 / 6**	**100**
	passer	1 / 2	50
	tomber	2 / 11	18,2
	sortir	2 / 5	40
Madeleine	***déménager***	**6 / 6**	**100**
	tomber	1 / 4	25
	descendre	**1 / 1**	**100**
	sortir	1 / 9	11,1
Nathan	***déménager***	**2 / 2**	**100**
	passer	**1 / 1**	**100**
	tomber	**3 / 3**	**100**
Virginie	***passer***	**1 / 1**	**100**
	sortir	1 / 2	50
	aller	1 / 14	7,1
Marc	***passer***	**1 / 1**	**100**
	monter	1 / 4	25
	tomber	2 / 4	50
	rester	1 / 3	33,3
Annie	*passer*	3 / 6	50
	rentrer	1 / 2	50
Kim	*déménager*	2 / 4	50
	tomber	**2 / 2**	**100**
	sortir	**1 / 1**	**100**
Jessica	*déménager*	2 / 3	66,7
	rester	**1 / 1**	**100**
Marie-Jeanne	*déménager*	4 / 5	80
Martin	***monter***	**2 / 2**	**100**

	tomber	1 / 5	20
	rentrer	**1 / 1**	**100**
	partir	1 / 4	25
Johanne	***passer***	**1 / 1**	**100**
	tomber	2 / 6	33,3
	descendre	1 / 2	50
	retourner	1 / 4	25
Martine	***monter***	**1 / 1**	**100**
	tomber	**1 / 1**	**100**
	rentrer	3 / 10	30
Sabrina	***déménager***	**3 / 3**	**100**
	sortir	1 / 2	50
Caroline	***déménager***	**1 / 1**	**100**
	monter	1 / 2	50
	tomber	1 / 2	50
	rentrer	4 / 8	50
	rester	2 / 9	22,2
	partir	2 / 25	8
Justin	*sortir*	1 / 2	50
	rester	**1 / 1**	**100**
	aller	1 / 17	5,9
Denise	***déménager***	**1 / 1**	**100**
	rester	**1 / 1**	**100**
Guylaine	*déménager*	4 / 5	80
	sortir	1 / 3	33,3
Marie-Laurence	*tomber*	1 / 4	25
Annouck	***passer***	**3 / 3**	**100**
	rentrer	1 / 4	25
Charles-Antoine	***tomber***	**1 / 1**	**100**
	rester	2 / 7	28,6
Mathieu	***passer***	**1 / 1**	**100**
	venir	1 / 5	20
David	***passer***	**1 / 1**	**100**
	monter	**1 / 1**	**100**
	tomber	1 / 2	50
	descendre	**1 / 1**	**100**
Florence	*déménager*	1 / 2	50

	passer	**1 / 1**	**100**
Hugo	*venir*	1 / 2	50
Denis	*monter*	2 / 6	33,3
	passer	1 / 5	20
	tomber	1 / 2	50
Richard	*partir*	1 / 3	33,3
Paul	*tomber*	2 / 6	33,3
Carl	*tomber*	1 / 2	50
Amélie	***déménager***	**2 / 2**	**100**
Christine	*sortir*	1 / 2	50
Mario	*passer*	1 / 2	50

Table 9.2 *Avoir* tokens by speaker (in decreasing order of *avoir* selection) and by lexical item (in decreasing order of overall *avoir* selection in the corpus).

9.7. Appendix 4C: Pronominal verbs that surfaced in the Montréal French corpus (ranked in decreasing order of *avoir* selection)

Non-standard uses have also been specified: a single asterisk (*) indicates that the verb is a Québecism and a double asterisk (**) indicates moreover that the use of that verb is also an Anglicism. In the cases where some verbs have more than one translation, the translations chosen here are the ones that best illustrate the use of the verb, as it appeared most frequently in the corpus.

Pronominal verbs	**N of *avoir* tokens**	**N of tokens**	**% *avoir***
se douter 'to doubt'	1	1	100
se déguiser 'to dress up'	1	1	100
s'habituer 'to get used to'	1	1	100
se tenir 'to hold onto'	2	4	50
se fouler 'to strain'	1	2	50
se marier 'to get married'	1	4	25
s'ouvrir 'to open'	1	4	25
s'arranger 'to sort itself out'	1	10	10
**se péter* 'to break X'	1	16	6,3
se coucher 'to go to bed'	1	28	3,6
se rendre (compte) 'to realise'	1	39	2,6
se faire + V 'to have X happen to o.s.'	2	94	2,1
se passer 'to go (well/ badly)'	2	98	2
se lever 'to get up'	1	59	1,7
se dire 'to say to oneself or each other'	0	118	0
se faire	0	66	0

'to do something for o.s.'			
se rencontrer 'to meet (each other)'	0	41	0
se mettre à + V 'to start doing X'	0	30	0
se réveiller 'to wake up'	0	30	0
se retrouver 'to end up'	0	27	0
se voir 'to see o.s. or each other'	0	26	0
se casser 'to break X'	0	23	0
se ramasser 'to end up'	0	21	0
se rendre 'to go to"	0	18	0
se promener 'to go for a walk'	0	17	0
s'acheter 'to buy X for o.s.'	0	13	0
se trouver 'to find o.s.'	0	11	0
s'inscrire 'to register'	0	10	0
se mettre 'to go' or 'to put on'	0	10	0
se parler 'to talk to each other'	0	9	0
se prendre 'to catch' or 'to order for o.s.'	0	9	0
se relever 'to get up'	0	9	0
s'endormir 'to fall asleep'	0	8	0
se blesser 'to hurt o.s.'	0	8	0
se brûler 'to burn o.s.'	0	8	0

s'installer 'to set up'	0	7	0
se baigner 'to go swimming'	0	7	0
se connaître 'to get to know'	0	7	0
**se pogner* 'to get', 'to hook up with someone' or 'to have an argument with someone'	0	7	0
s'asseoir 'to sit down'	0	6	0
se faire (mal) 'to hurt o.s.'	0	6	0
se laisser 'to let o.s. X' or 'to break up'	0	6	0
se revoir 'to see each other again'	0	6	0
s'en aller 'to leave'	0	5	0
s'habiller 'to get dressed'	0	5	0
se cogner 'to bump into'	0	5	0
se croiser 'to pass each other'	0	5	0
se donner 'to give o.s. or each other'	0	5	0
se lancer 'to throw to each other' or 'to embark'	0	5	0
se perdre 'to get lost'	0	5	0
se préparer 'to get ready'	0	5	0
se rappeler 'to remember'	0	5	0
se sortir 'to extricate o.s.'	0	5	0
**se tanner*	0	5	0

'to get tired of'			
s'améliorer 'to improve'	0	4	0
s'apercevoir 'to realise'	0	4	0
s'enfuir 'to run away'	0	4	0
se brosser 'to brush'	0	4	0
se fréquenter 'to date'	0	4	0
se laver 'to wash o.s.'	0	4	0
se planter 'to mess up'	0	4	0
se séparer 'to separate'	0	4	0
s'appeler 'to call each other'	0	3	0
s'arrêter 'to stop'	0	3	0
s'avancer 'to move forward'	0	3	0
**s'en venir* 'to arrive'	0	3	0
s'impliquer 'to get involved'	0	3	0
s'occuper 'to deal with'	0	3	0
se briser 'to break'	0	3	0
**se chicaner* 'to argue'	0	3	0
se comprendre 'to understand'	0	3	0
se débrouiller 'to manage'	0	3	0
se déchirer 'to rip'	0	3	0

se dépêcher 'to hurry'	0	3	0
**se distancer* 'to grow distant'	0	3	0
se fermer 'to close'	0	3	0
se foncer 'to charge into'	0	3	0
se forcer 'to make an effort'	0	3	0
se libérer 'to free o.s.'	0	3	0
se payer 'to treat o.s.'	0	3	0
se pointer 'to turn up'	0	3	0
se poser 'to land' or 'to come up'	0	3	0
se présenter 'to introduce o.s.'	0	3	0
se recoucher 'to go back to bed'	0	3	0
se remettre 'to get better'	0	3	0
se remettre à + V 'to start doing X again'	0	3	0
se retirer 'to remove o.s.'	0	3	0
se retourner 'to turn around'	0	3	0
se sentir 'to feel'	0	3	0
se souvenir 'to remember'	0	3	0
se taper 'to hit o.s. or each other'	0	3	0
se tromper 'to make a mistake'	0	3	0
s'accrocher	0	2	0

'to hang on'			
s'affirmer 'to assert o.s.'	0	2	0
s'amuser to have fun'	0	2	0
s'écraser 'to crash'	0	2	0
s'embrasser 'to kiss'	0	2	0
**s'enfarger* 'to stumble'	0	2	0
s'entraîner 'to train'	0	2	0
**s'évacher* 'to slump down'	0	2	0
s'infiltrer 'to infiltrate'	0	2	0
se changer 'to get changed'	0	2	0
se coincer 'to get stuck'	0	2	0
se construire 'to build'	0	2	0
se fendre 'to crack'	0	2	0
se fiancer 'to get engaged'	0	2	0
se frapper 'to hit o.s. or each other'	0	2	0
se louer 'to rent'	0	2	0
se poursuivre 'to carry on'	0	2	0
se produire 'to happen'	0	2	0
se rajouter 'to add'	0	2	0
se réchauffer 'to warm up'	0	2	0

se rendormir 'to fall back asleep'	0	2	0
se renseigner 'to find out'	0	2	0
se rentrer 'to crash into'	0	2	0
se répandre 'to spill'	0	2	0
se stationner 'to park'	0	2	0
se surprendre 'to find o.s. doing X'	0	2	0
se tuer 'to kill o.s.'	0	2	0
s'adapter 'to adapt o.s.'	0	1	0
s'allumer 'to light'	0	1	0
s'attacher 'to fasten up'	0	1	0
**s'auto-donner* 'to give o.s. X'	0	1	0
s'échanger 'to exchange'	0	1	0
s'échapper 'to escape'	0	1	0
s'échoir 'to fall due'	0	1	0
s'éclater 'to explode'	0	1	0
s'effectuer 'to make', 'to happen' or 'to take place'	0	1	0
**s'effoirer* 'to slump down'	0	1	0
s'éloigner 'to move away'	0	1	0
s'emmagasiner 'to store up'	0	1	0
s'endetter	0	1	0

'to get into debt'			
s'enfermer 'to shut o.s. in'	0	1	0
s'engueuler 'to argue'	0	1	0
s'enregistrer 'to record o.s.'	0	1	0
s'entendre 'to come to an agreement'	0	1	0
s'équiper 'to kit o.s. out'	0	1	0
s'étirer 'to stretch'	0	1	0
s'excuser 'to apologise'	0	1	0
s'inquiéter 'to worry'	0	1	0
s'isoler 'to isolate o.s.'	0	1	0
s'obstiner 'to persist in doing'	0	1	0
se badigeonner 'to smear'	0	1	0
se barrer 'to lock'	0	1	0
se baser 'to base'	0	1	0
se battre 'to fight'	0	1	0
se bloquer 'to lock'	0	1	0
***se booker* 'to book'	0	1	0
***se bumper* 'to bump'	0	1	0
se cacher 'to hide o.s.'	0	1	0
se calmer 'to calm o.s.'	0	1	0
se camoufler	0	1	0

'to camouflage o.s.)			
***se checker* 'to check out o.s. or each other'	0	1	0
se choisir 'to choose o.s.'	0	1	0
se claquer 'to pull (a muscle)' or *'to wolf down'	0	1	0
**se coller* 'to cuddle'	0	1	0
se commander 'to order in'	0	1	0
se continuer 'to continue'	0	1	0
se côtoyer 'to come close'	0	1	0
se couvrir 'to cover up'	0	1	0
se craquer 'to crack'	0	1	0
**se crisser* 'to not give a damn'	0	1	0
se culminer 'to reach a peak'	0	1	0
***se dater* 'to date'	0	1	0
se débarrasser 'to get rid of'	0	1	0
se décider 'to make a decision'	0	1	0
se découvrir 'to discover o.s. or each other'	0	1	0
se demander 'to ask o.s. or each other'	0	1	0
se dépeigner 'to ruffle oneself's hair'	0	1	0
se déplacer 'to move'	0	1	0
se déployer 'to spread'	0	1	0

se déposer 'to settle'	0	1	0
se disloquer 'to dislocate'	0	1	0
se dissiper 'to drift away'	0	1	0
se familiariser 'to get familiar with'	0	1	0
se fier 'to trust'	0	1	0
se finir 'to finish'	0	1	0
se fixer 'to set o.s. X'	0	1	0
**se flexer* 'to leave'	0	1	0
**se fourrer* 'to make a mistake'	0	1	0
se foutre 'to not give a damn'	0	1	0
se fracasser 'to crash against X'	0	1	0
se fracturer 'to fracture'	0	1	0
se frustrer 'to get frustrated'	0	1	0
se gangréner 'to gangrene'	0	1	0
**se garrocher* 'to throw o.s. at X'	0	1	0
**se jaser* 'to chatter'	0	1	0
se jeter 'to throw o.s.'	0	1	0
se lier 'to strike up friendship with'	0	1	0
***se lighter* 'to light'	0	1	0
se manquer	0	1	0

'to miss'			
***se minder* 'to mind o.s. to do X'	0	1	0
se monter 'to get worked up'	0	1	0
se normaliser 'to normalise'	0	1	0
***se parker* 'to park'	0	1	0
se peigner 'to comb'	0	1	0
se perpétuer 'to be perpetuated'	0	1	0
se placer 'to place'	0	1	0
se plier 'to submit to'	0	1	0
se pratiquer 'to practice'	0	1	0
se prêter 'to lend itself to X'	0	1	0
se prévoir 'to plan'	0	1	0
se procurer 'to procure'	0	1	0
se prolonger 'to last'	0	1	0
se racheter 'to buy again for o.s.' or 'to make it up to'	0	1	0
se ramener 'to come down to'	0	1	0
se rapporter 'to refer to' or 'to bring back for o.s.'	0	1	0
se rapprocher 'to get close to'	0	1	0
se rasseoir 'to sit back down'	0	1	0
se rattraper 'to redeem o.s.'	0	1	0

se rebloquer 'to lock again'	0	1	0
se réconcilier 'to make up'	0	1	0
se réessayer 'to try again'	0	1	0
se refaire 'to make again'	0	1	0
se réinscrire 'to register again'	0	1	0
se relocaliser 'to relocate'	0	1	0
se reparler 'to talk to each other again'	0	1	0
se reperdre 'to get lost again'	0	1	0
se reprendre 'to correct o.s.'	0	1	0
se réserver 'to set aside for o.s.'	0	1	0
se retenir 'to stop o.s. from'	0	1	0
***se scrapper* 'to ruin'	0	1	0
se servir 'to help o.s.'	0	1	0
se soigner 'to treat o.s.'	0	1	0
**se sonner* 'to shock'	0	1	0
se vomir 'to vomit on o.s.'	0	1	0

Table 9.3 Pronominal verbs that surfaced in the Montréal French corpus (ranked in decreasing order of *avoir* selection)