

Viewing Romance through a variationist lens

Shana Poplack
University of Ottawa
shana.poplack@uottawa.ca



How to cite: Poplack, Shana. 2023. Viewing Romance through a variationist lens. In *Romance grammars: context and contact*, eds. Alice Corr & Norma Schifano. *Special issue of Isogloss. Open Journal of Romance Linguistics* 9(2)/10, 1-11.
DOI: <https://doi.org/10.5565/rev/isogloss.298>

1. Viewing Romance through a variationist lens

As a sociolinguist, my work is informed by the dual mandate of studying language in its societal context (i.e. as it is actually spoken in some well-defined speech community) and privileging linguistic issues that concern that community. Now the minute you pay even a modicum of attention to language as it's spoken, you come face to face first, with its core property—inherent variability—and then, with the decision of how to handle it. I choose to deal with variability rather than ignore it, and this is where the specifically variationist perspective comes in. In my own case, it involves not just a casual interest in linguistic variables, but rather, a serious commitment to a scientific framework involving accountable, replicable, systematic quantitative research on actual language *use*. In this context, the Romance angle proceeds from the societal: when I worked in New York City, with its huge Hispanic population, the main focus was on Spanish; in Canada, it is French that is foregrounded. For many years now I've run a Sociolinguistics Laboratory (www.sociolinguistics.uottawa.ca) at the University of Ottawa whose work combines all of these elements. In what follows, I briefly illustrate how we go about this.

2. Data

Appropriate data are of prime importance to the variationist. In the first instance, they should be representative, of both the speech community under investigation and of the way its members actually speak. The latter imperative explains why we tend to target the *vernacular*, characterized as the most regular and systematic form of language (Labov [1966] 2006). They should also be appropriate to the research question under consideration. I noted above that many of our Lab’s projects are inspired by society’s take on linguistic matters. As but one example, part of the received wisdom about Canadian French is that it has been deteriorating since its split from its European source, and contact with English, the majority language in most of the country, is thought to be a major driver. This state of affairs is exacerbated by the widespread belief that French language-arts teachers don’t speak French well enough to teach it. So the major *linguistic* preoccupations of Canadian francophones involve assimilation, attrition and the resulting “quality” of French. Because these deal with change, they are diachronic in nature. (If the quality of French is poor today, it must have been better before. If the language has deteriorated, from what? And so on.) Since change cannot be studied without reference to an earlier stage, we built a number of large-scale corpora specifically designed to address these issues empirically. Corpus creation and data handling at the Sociolinguistics Lab are detailed in Poplack (2022); here we illustrate with a few that are particularly relevant to the present discussion.

The Ottawa-Hull corpus (Poplack 1989), a massive bilingual dataset collected in the 1980s from a random sample of francophones residing in the national capital region of Canada, is stratified in several ways, but especially according to minority vs majority status of French. About 25 years thereafter we returned to one of the areas sampled previously and recorded francophone high-schoolers and their French language arts teachers (Poplack 2015). Along the way, in an incredible stroke of luck, we discovered a cache of audio recordings of rural Québécois made by folklorists in the 1940s and 50s, but born as far back as 1846! So we built a corpus out of those (Poplack & St-Amand 2007). Together these three datasets cover an apparent-time span of nearly a century and a half (from the dates of birth of the oldest to the youngest speakers)—61 years in real time. Either way, a nearly unprecedented period over which to chart any changes to the structure of the spoken language, as well as to study the persistence, evolution and effects of language mixing involving French. As depicted in Table 1, these corpora include over 4 million words of spontaneous speech, more than enough to make meaningful measures of frequency. And the hundreds of speakers analyzed enable us to measure diffusion, both over time and across it. All of these data have been transcribed, computerized and concordanced, for the purpose of large-scale quantitative study.

Table 1. A century and a half of spoken French

Century represented	N speakers	N words
19 th	37	524,900
20 th	120	2,500,000
21 st	166	1,139,766
Overall	323	4,163,776

3. Variability

I noted above that spontaneous speech means inherent variability. Some notable examples, taken from French morphosyntax, include the alternation between: the negative particle *ne* and a null variant \emptyset in the negative utterances illustrated in (1), the 1st p. pl. pronoun *nous* ‘we’ and the 3rd p. sg. impersonal *on* ‘one, they’ in 1st person plural inclusive contexts in (2), subjunctive, conditional and indicative variants under “subjunctive-selecting” governors (3) and the synthetic, periphrastic and futurate present variants in future temporal reference contexts (4).¹

- (1)a. Il *n*'y a rien à faire. (21C.546.547)
 ‘There’s nothing to do.’
 b. Parce que les salaires [\emptyset] valent pas l’effort que tu mets. (20C.061.73)
 ‘Because the salaries aren’t worth the effort you put in.’
- (2) Alors depuis euh- à peu près cinq, six ans, *nous*_[1PL] suivons *nos* élèves. Les professeurs de français et de mathématiques, *on*_[3SG] suit *nos* élèves deux ans. (21C.317.549)
 ‘So for about five, six years, we’ve followed our students. Us French and math teachers, we follow our students for two years.’
- (3)a. Tu aimerais mieux qu’ils *soient*_[SBJV] pas là. (20C.25.608)
 ‘You’d like it better if they weren’t there.’
 b. Tu aimerais qu’il *guérisse*_[IND] ta paralysie. (20C.044.1554)
 ‘You’d like it if he cured your paralysis.’
 c. J’aimerais ça que vous *verriez*_[COND] mon mari. (20C.072.1335)
 ‘I’d like it if you could see my husband.’
- (4)a. Mais le français c’est sûr ça *arrivera*_[SF] pas demain là. (21C.306.405)
 ‘But French, for sure it *won’t* be for tomorrow.’
 b. Là il y en a une des blessées qui *va revenir*_[PF] là comme demain. (21C.109.295)
 ‘One of the wounded is *going to return* like tomorrow.’
 c. Fait que là je *commence*_[P] demain. (21C.150.16)
 ‘So I *start* tomorrow.’

Why all this variability? Those who subscribe to the myth of form-function symmetry, which holds that to every difference in form corresponds a difference in meaning, construe it as change. In the Canadian case, the belief is that change has taken place on Canadian soil, i.e. in the transplanted dialect. Confirmation of such a scenario requires going much further back in time, before the colonization of New France. This prompted us to assemble yet another corpus (surely the weirdest in our lab), the *Recueil Historique des grammaires du français* (RHGF; Poplack et al. 2015). This is a compilation of 163 prescriptive grammars dating from 1530 to 1999. The

¹ In the present article, codes in parentheses refer to corpus (19C = *Récits du français québécois d’autrefois* [Poplack & St-Amand 2007], 20C = *Corpus du français parlé à Ottawa-Hull* [Poplack 1989], 21C = *Français en contexte, milieux scolaire et social* [Poplack 2015]), speaker and line number. Examples are reproduced verbatim from audio recordings.

logic behind looking to prescriptive grammars as an earlier stage of speech is that if grammarians were invoking certain forms (usually to condemn them), we can be pretty sure that someone was already using them. One startling finding of this project is that the competing variants, of these and the many other variables we've studied, were all attested from the earliest times—a nice illustration of the ubiquity of inherent variability, as well as a reminder that it does not in and of itself constitute evidence of change. Variability is a necessary precursor to change, granted, but should not be confounded with it. Variant rates, and especially the *conditioning* of variant choice, are far more revealing.

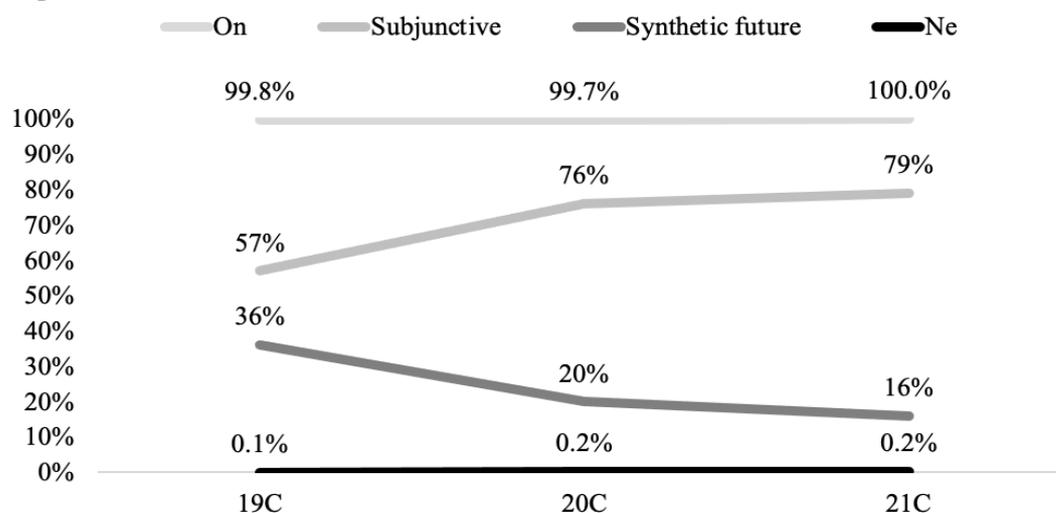
4. The conditioning of variability

What do we learn from these? In accordance with the *Principle of Accountability* (Labov 1966/2006; 1972), which enjoins us to contextualize the variant of interest with respect to every context in which it could have occurred, *even if it did not*, we systematically search the corpora to find every example of a *variable context*, defined as the locus in discourse where variants may alternate with no change in referential meaning, as in examples (1)-(4) above. We can then systematically record the variant that was selected in each. This tends to result in massive quantitative analyses, as depicted in Table 2, enhancing the confidence we can have in the results.

Table 2. Tokens analyzed

Variable	19C	20C	21C	Total
Future temporal reference	4,691	3,559	3,656	11,906
Negation	9,438	61,316	14,693	85,447
1 st person plural	3,292	25,457	8,892	37,641
Expression of the subjunctive	583	2,767	1,090	4,440
Total N tokens	18,004	93,099	28,331	139,434

Calculation of rates of variant selection over the century and a half represented by our corpora (Figure 1) suggests that two of the variables we have been illustrating with have remained stable: *on* has prevailed over *nous* since the 19th century (at least), and has now succeeded in ousting it altogether, while *ne* has almost always been absent from negative sentences over the same period. If these are changes, they have apparently gone to completion. Meanwhile, selection of the subjunctive variant has risen substantially, while the synthetic future has declined by about the same amount.

Figure 1. Variant rates over time

Now rate changes are the first (and often the only) line of evidence linguists marshal to infer change, and rightfully so. But they're not foolproof, because they may fluctuate according to the vagaries of the situation. A hyperformal variant may appear to have waned in corpora made up of vernacular speech, for example. The *conditioning*, or the configuration of constraints promoting the selection of one linguistic variant of a linguistic variable over another, is both more stable and more revealing. Accordingly, we construe variable conditioning as a portion of the grammar underlying the variation.

By way of illustration, consideration of the factors contributing to choice of the synthetic future [SF] (Poplack & Dion 2009; Poplack & Turpin 1999) teaches us that this variant is pretty much restricted to negated utterances. This is illustrated in (5), where the affirmative future predications are all expressed with the periphrastic future [PF]; only their negative counterpart features SF.

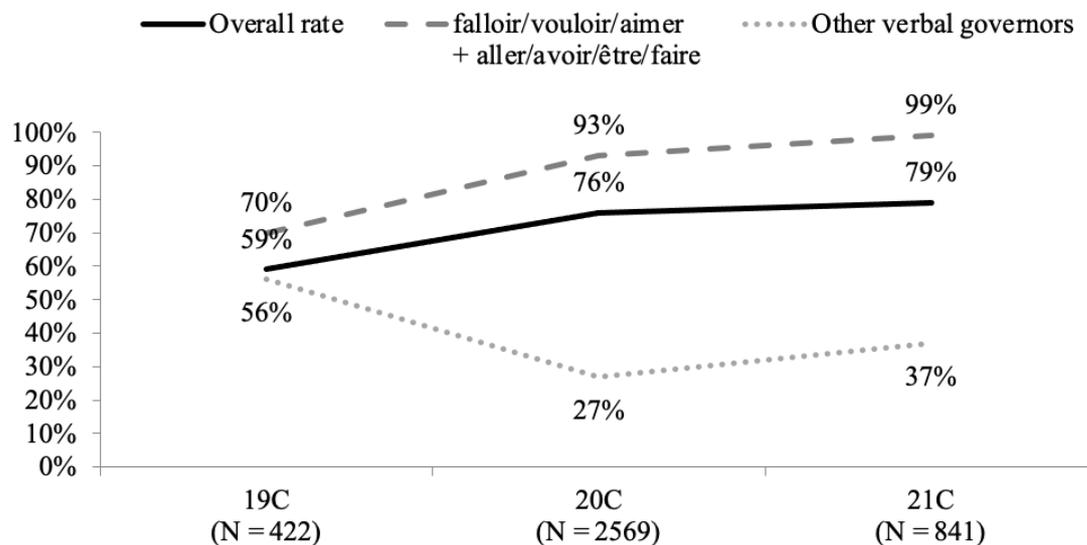
- (5) ...il *va virer*_[PF] la clef, il *va* la *rebarrer*_[PF] puis *va* la *redébarrer*_[PF]. Il *s'apercevro*_[SF] pas qu'elle était débarrée. (19C.021.1256)
 '...he's *going to turn* the key, he's *going to lock* it again, then *going to unlock* it again. He *won't* notice that it was unlocked.'

Likewise, the subjunctive variant turns out to be basically entrenched in a handful of embedded verbs under a handful of governors (Poplack 1992; Poplack, Lealess & Dion 2013), and both the negative particle *ne* and 1st p. pl. inclusive *nous* have assumed new functions (Poplack 2015; Poplack & Dion 2021; Poplack & St-Amand 2007). None of these uses is taught in school (on the contrary!), yet they all form part of the implicit usage norms that community members implement regularly, often unawares. Such structural details could not be intuited from casual inspection.

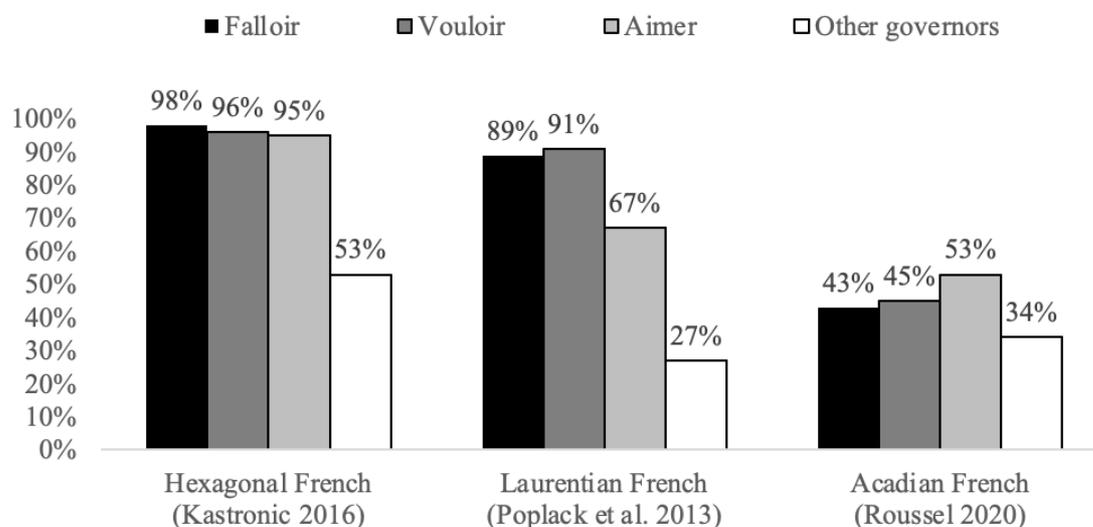
Armed with this information, we can achieve a much more nuanced view of actual linguistic developments. For example, we learn that the two variables that appear, from variant *rates*, to have remained stable across time (Figure 1), have in fact changed. Both *ne* and *nous* were resuscitated from the dead (quite literally, in the case of the latter) to serve a novel discourse function: that of hyperstyle marker. And the apparent increase in subjunctive selection actually signals a *decrease* in productivity,

because what changed is that the subjunctive variant is becoming more entrenched under just a few main, and to a lesser extent, embedded verbs, instead of being dispersed across a wide range of governors. Everywhere else, it has receded (Figure 2). Poplack & Dion (2021) outline a variety of changes in French morphosyntax on which rate differences are either silent or misleading.

Figure 2. Entrenchment of the subjunctive over time



Even more revealing, we can make use of these conditioning profiles as benchmarks for comparison, both synchronic and diachronic. For example, we can assess how a given form is used in what are traditionally viewed as substantively different varieties of the same languages. Drawing again upon the behaviour of the Laurentian French subjunctive described in Poplack (1992), Poplack, Lealess & Dion (2013) and above, Kastronic (2016) and Roussel (2020) compared the situation in Hexagonal and Acadian French respectively. Despite some inter-variety differences in overall *rate* of subjunctive selection, the similarities in conditioning are patent, as can be seen from the shared tendency toward entrenchment under the same cohort of governors (Figure 3). Such parallels, little short of remarkable in view of the widespread perception of differences amongst these varieties, had never been previously reported. They could not have been uncovered without first recognizing, and then systematically analyzing, the inherent variability of subjunctive selection in French.

Figure 3. Subjunctive rate according to governor in three varieties of French

5. Romance

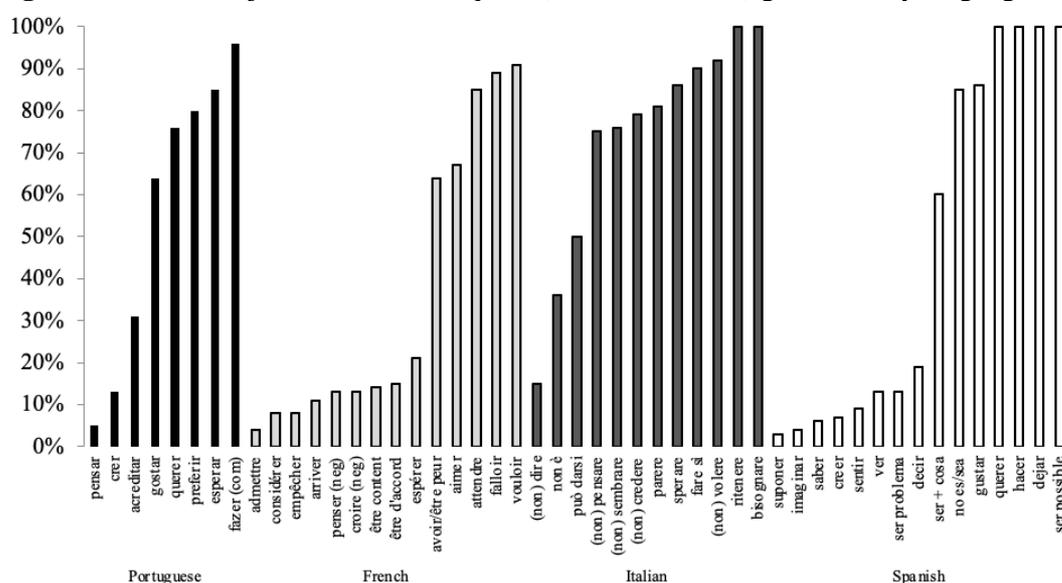
By the same token, we can extend the comparison *cross*-linguistically, as exemplified by a recent study of subjunctive use in four Romance languages (Poplack et al. 2018). From systematic inspection of corpora of spontaneous speech, we detected the same kind of variability (in subjunctive-selecting contexts, of course) reported for French in each of Portuguese, Italian and Spanish, even under the same governor, as illustrated with the verb *believe* in examples (6)-(9).²

- (6)a. Je crois pas que ce *soit*_[SBJV] la fin du monde. (FR.060.195)
 ‘I don’t believe that it would be the end of the world.’
 b. Je crois pas que l’âge *a*_[IND] tant à faire que ça. (FR.003.189)
 ‘I don’t believe that age has that much to do with it.’
- (7)a. Eu acredito que *vá*_[SBJV] *sair*. (PTG.143.356)
 ‘I believe that it will come out.’
 b. Eu acredito que ele *devia*_[IND] *ter* em torno de setenta anos de idade. (PTG.99.148)
 ‘I believe that he must be around seventy years old.’
- (8)a. Credevo *fosse*_[SBJV] un errore. (ITA.023.152)
 ‘I believed it was a mistake.’
 b. E tutti che dicevano- credevano che *ammazzava*_[IND] uomini. (ITA.304.10)
 ‘And everybody that said- believed that he killed men.’
- (9)a. No creo que *haya*_[SBJV] nadie aquí que no pague. (SPN.073.668)
 ‘I don’t believe that there is anybody here who doesn’t pay rent.’
 b. No creo que *hay*_[IND] que firmar. (SPN.086.555)
 ‘I don’t believe that you have to sign.’

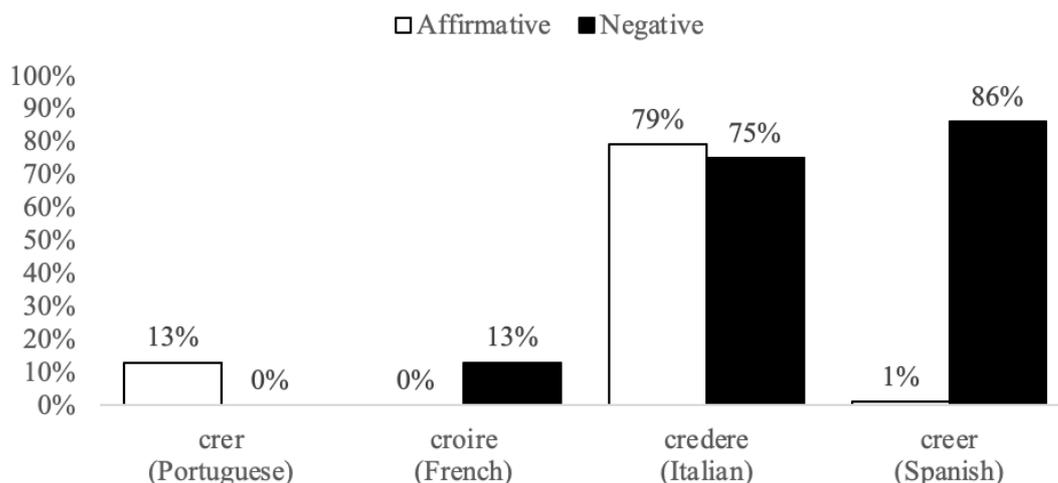
² Examples in (6)-(9) taken from Poplack et al. (2018).

Less predictably, however, overall rates of the subjunctive variant fluctuate wildly cross-linguistically, from a low of 37% in Spanish to a high of 76% in French (Poplack et al. 2018: 229). This is puzzling to those who maintain that the subjunctive is obligatory in each of these languages, whether to convey certain semantic readings, or because certain matrix verbs subcategorize for it. Our method also reveals, even more unexpectedly, that Spanish, the Romance language with the lowest rate of the subjunctive, is the one that actually uses it *most* productively. As noted above, the opposite is true of French, which boasts the highest rate. Moreover, Figure 4 shows that the strong lexical component of subjunctive selection observed for French (Figure 3) applies to the other Romance languages as well: a small cohort of governors is heavily associated with the subjunctive variant, while another is specifically disfavoured to it, with few in the middle.

Figure 4. Rate of subjunctive across frequent (10+ occurrences) governors by language



But again surprisingly, given their shared descent from the same Latin ancestor, the effect doesn't necessarily involve the same lexical verbs! This can be seen from the behaviour of *believe* in Figure 5. In Spanish, *believe* basically only selects subjunctive when negated; the same is true of French, even if to a far lesser extent. In Portuguese, on the other hand, the rare occasions where *believe* co-occurs with a subjunctive are in *affirmative* contexts, while in Italian, it is strongly favoured independent of polarity.

Figure 5. Rate of subjunctive for the verb *think* by polarity of the clause across languages

The foregoing are just a few examples of the surprising facts about language *use* we discover by analyzing variability. Other areas of French grammar we have studied include question formation (Elsig 2009; Elsig & Poplack 2006), null subjects (Leroux & Jarmasz 2006), possessive marking (Miller & Dion 2009; Poplack 2015), the expression of necessity (Kastronic 2016), complementizer realization (Dion 2003), hypothetical *si*-clauses (Poplack 2015) and auxiliary alternation (Willis 2000). We like to think that these are of theoretical interest in their own right—perhaps best illustrated by the robust finding that in spontaneous speech, the Romance subjunctive has little if any semantic component, in striking contrast to the massive literature devoted to elucidating its meaning. Once we have obtained this kind of structural information, we can marshal it as the basis for diachronic and synchronic comparisons, like those illustrated above. Viewed through a variationist lens, the Romance languages thus offer a compelling conduit into a wide variety of linguistic and extra-linguistic issues.

Acknowledgments

The research on which these remarks are based was generously funded by the Social Sciences and Humanities Research Council of Canada in the form of a Canada Research Chair in Linguistics (2001-2022) and numerous research grants. I am grateful to the many Sociolinguistics Lab associates who contributed to this work, and to Nathalie Dion, who coordinated all their efforts.

References

- Dion, Nathalie. 2003. *L'effacement du que en français canadien : Une étude en temps réel*. M.A. mémoire, University of Ottawa.
- Elsig, Martin. 2009. *Grammatical Variation Across Space and Time: The French Interrogative System*. Amsterdam/Philadelphia: John Benjamins Publishing.

-
- Elsig, Martin, & Shana Poplack. 2006. Transplanted dialects and language change: Question formation in Québec. *University of Pennsylvania Working Papers in Linguistics*. 12(2): 77-90.
- Kastronic, Laura. 2016. *A comparative variationist approach to morphosyntactic variation in contemporary Hexagonal and Quebec French*. Ph.D. dissertation, University of Ottawa.
- Labov, William. 1966/2006. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Leroux, Martine, & Lidia-Gabriela Jarmasz. 2006. A study about nothing: Null subjects as a diagnostic of convergence between English and French. *University of Pennsylvania Working Papers in Linguistics* 12(2): 1-14.
- Miller, Kim, & Nathalie Dion. 2009. Measuring the effectiveness of proscription: Possessive prepositions in French. Paper presented at NWA V 38. University of Ottawa.
- Poplack, Shana. 1989. The care and handling of a mega-corpus. In Ralph W. Fasold, & Deborah Schiffrin (eds.), *Language Change and Variation*, 411-451. Amsterdam: Benjamins.
- Poplack, Shana. 1992. The inherent variability of the French subjunctive. In Christiane Laeuffer, & Terrell A. Morgan (eds.), *Theoretical Analyses in Romance Linguistics*, 235-263. Amsterdam: Benjamins.
- Poplack, Shana. 2015. Norme prescriptive, norme communautaire et variation diaphasique. In Kirsten Kragh & Jan Lindschouw (eds.), *Variations diasystématiques et leurs interdépendances*, 293-319. Série TraLiLo, Strasbourg: Société de linguistique romane.
- Poplack, Shana. 2020. A variationist perspective on language contact. In Evangelia Adamou, & Yaron Matras (eds.), *Routledge Handbook of Language Contact*, 46-62. New York & London: Routledge.
- Poplack, Shana, & Nathalie Dion. 2009. Prescription vs praxis: The evolution of future temporal reference in French. *Language* 85(3): 557-587.
- Poplack, Shana, & Nathalie Dion. 2021. Cartographie de la variation et du changement morphosyntaxique en français : leçons à retenir. *Cahiers Internationaux de Sociolinguistique* 18: 81-113.
- Poplack, Shana, Lidia-Gabriela Jarmasz, Nathalie Dion, & Nicole Rosen. 2015. Searching for “Standard French”: The construction and mining of the *Recueil historique des grammaires du français*. *Journal of Historical Sociolinguistics* 1(1): 13-56.

Poplack, Shana, Allison Lealess, & Nathalie Dion. 2013. The evolving grammar of the French subjunctive. *Probus* 25(1) (Special 25th anniversary issue), 139-195.

Poplack, Shana, & Anne St-Amand. 2007. A real-time window on 19th century vernacular French: The *Récits du français québécois d'autrefois*. *Language in Society* 36(5), 707-734.

Poplack, Shana, & Danielle Turpin. 1999. Does the FUTUR have a future in (Canadian) French? *Probus* 11(1): 133-164.

Poplack, Shana, Rena Torres Cacoullos, Nathalie Dion, Rosane Berlinck, Salvatore Digesto, Dora LaCasse, & Jonathan Steuck. 2018. Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. In Wendy Ayres-Bennett & Janet Carruthers (eds.), *Manuals in Linguistics: Romance Sociolinguistics*, 217-252. Berlin/Boston: de Gruyter.

Roussel, Basile. 2020. *À la recherche du temps (et des modes) perdu(s) : Une étude variationniste en temps réel du français acadien parlé dans le nord-est du Nouveau-Brunswick*. Ph.D. dissertation, University of Ottawa.

Willis, Lauren. 2000. *Être ou ne plus être: Auxiliary alternation in Ottawa-Hull French*. Ph.D. dissertation, University of Ottawa.