CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

The acquisition of variation and change in English by Canadian francophones

Stephen Levey[1], Laura Kastronic[1] and Yasmine Abou Taha[2]

[1]University of Ottawa, Atlantic Canada Opportunities Agency, and [2]University of Toronto
Corresponding author: Stephen Levey; Email: slevey@uOttawa.ca

Abstract

This article describes the compilation of three new spoken language corpora designed to address the acquisition of variation and change in English by second language (L2) learners. The first corpus represents L2 English recorded from native Canadian francophones in the Canadian National Capital Region. A second corpus of vernacular English recorded from native anglophones in the same region furnishes a local baseline variety of the target language (TL). A third corpus of local Canadian French represents L2 speakers' first language (L1). These corpora are used to determine the extent to which L2 variable patterns approximate those in the TL. Comparison of L2 variable features with structural analogues in speakers' L1, French, additionally affords a window on possible L1 transfer effects. We present a case study exploring the L2 acquisition of quotative variation and change in the local TL benchmark variety. Results point to the close approximation of higher-proficiency L2 speakers to TL community norms, challenging the prevailing notion that L2 acquisition typically involves incomplete mastery of TL patterns and constraints.

Keywords: L2 acquisition; comparative sociolinguistics; quotative variation

Résumé

Cet article traite de la compilation de trois nouveaux corpus de discours spontané pour aborder la manière dont les locuteurs de l'anglais langue seconde (L2) acquièrent la variation et le changement linguistique retrouvés dans le discours des locuteurs natifs de l'anglais. Le premier corpus représente l'anglais L2 employé par des francophones canadiens natifs enregistrés dans la région de la capitale nationale du Canada. Un deuxième corpus d'anglais vernaculaire enregistré auprès d'anglophones natifs dans la même région fournit un point de repère représentant la langue cible (LC). Un troisième corpus de français canadien enregistré dans la même localité représente la première langue (L1) des locuteurs L2. Ces corpus sont utilisés afin de préciser la mesure dans laquelle les patrons variables employés par les locuteurs L2 se rapprochent de ceux retrouvés dans la LC. La comparaison

de structures variables de la L2 avec des homologues structurels retrouvés dans la L1, le français, permet de mieux comprendre d'éventuels effets translinguistiques de L1 en L2. Par l'entremise d'une étude de cas, nous explorons l'acquisition L2 de la variation et du changement linguistique dans le système des formes d'introduction de discours direct employées par les locuteurs locaux de la LC. Les résultats indiquent que les apprenants avancés de l'anglais se rapprochent étroitement des normes linguistiques propres à la LC, remettant en question l'idée dominante que l'acquisition d'une L2 s'aboutit généralement à la maîtrise incomplète des patrons et des contraintes linguistiques retrouvés dans la LC.

1. Introduction

The last three decades have witnessed intensified interest in the investigation of second-language acquisition from a sociolinguistic perspective, encompassing diverse populations situated in varied learning environments, ranging from study-abroad contexts (e.g., Regan et al. 2009) and second-language immersion programs (e.g., Mougeon et al. 2010) to community-based settings (e.g., Sankoff et al. 1997, Blondeau et al. 2002). Multiple studies stress that attainment of native-like proficiency in a second language (L2) requires the learner to approximate the variable sociolinguistic patterns characteristic of the target language (TL) being acquired, in addition to mastering its relatively categorical features (see e.g. Rehner et al. 2003, Bayley and Regan 2004, Mougeon et al. 2004, Bayley and Preston 2008, Mougeon et al. 2010, Howard et al. 2013, Bayley and Escalante 2022). There is general agreement that for L2 speakers to approximate TL patterns of variation, they must not only employ the same repertoire of variants used by TL speakers, as well as reproducing their attendant usage rates, but that their speech should also reflect the operation of the same social and linguistic constraints governing variant selection in the community-based TL variety to which learners are routinely exposed (see e.g. Meyerhoff and Schleef 2014, Rehner and Mougeon 2022).

Within this research context, increasing attention has been paid to the L2 acquisition of TL *vernacular* norms, characteristically employed in unreflecting discourse when speakers are not engaged in monitoring their speech (Labov 1984). This line of inquiry has prompted a number of key questions. To what extent are those norms (in)completely acquired and how can this be assessed (see e.g. Howard et al. 2013)? How proficient must learners be in their L2 to master TL vernacular features? What is the role, if any, of L2 speakers' first language (L1) in mediating the acquisition of TL variables, especially in cases where the variables in question have structural counterparts in learners' L1 (see e.g. Odlin 2003)?

In Canada, cumulative progress in elucidating these questions has benefited extensively from research investigating the acquisition of French as a second language by Canadian anglophones, targeting a heterogeneous set of variable TL features (see e.g. Nagy et al. 1996, Sankoff et al. 1997, Nagy et al. 2003, Blondeau et al. 2014). But, remarkably, this line of research has paid correspondingly much less attention to the inverse state of affairs, namely, the acquisition of English by Canadian francophones.

In this article, we introduce a new project specifically designed to address this imbalance by focusing on the L2 acquisition of English by Canadian francophones located in the Canadian National Capital Region, the locus of extensive bilingualism and language contact (Poplack 2018). Integrating insights from variation theory (e.g., Labov 1972) with the methodological and analytical framework of comparative variationist sociolinguistics (e.g., Poplack and Meechan 1998, Poplack and Tagliamonte 2001), our investigation draws on a corpus of spontaneous L2 English recorded from Canadian francophones who acquired Canadian French as their primary language in childhood. A second corpus of vernacular English recorded from native anglophones in the Canadian National Capital Region furnishes a local benchmark variety of the TL. Innovative in this investigation is our use of a third corpus of local Canadian French obtained from a subset of the L2 speakers we recorded, who were interviewed in their native language in addition to English. Evidence abstracted from these three data sources enables us to (i) identify and characterize patterns of variation and change in the TL; (ii) effect systematic comparisons between the L2 corpus and the local TL benchmark variety, with a view to determining the extent to which L2 usage patterns approximate those in the TL; and (iii) compare L2 variable features with structural analogues in speakers' L1, French, in order to assess potential L1 influence on speakers' L2.

In the following sections, we first situate the project in the wider context of sociolinguistic research on L2 acquisition. We then describe the community targeted for investigation as well as the L2 population sample that supplied the data for analysis. We also briefly explore local attitudes towards acquiring and using English as a second language in view of the recognition that different affective orientations towards second-language learning are reported to have a bearing on levels of L2 attainment (Gardner 1985). We subsequently describe our methods of corpus compilation. Finally, we illustrate the utility of our materials for investigating second-language learning by presenting a case study exploring the L2 acquisition of ongoing change involving variable quotative expressions in the TL baseline variety.

2. Research context

There has been a longstanding preoccupation with variation in the study of L2 acquisition. Most early work focuses on the alternation of non-target and target forms in L2 acquirers' interlanguage (Selinker 1972) or developing grammar. This type of alternation is typically referred to as Type I variation (see Mougeon et al. 2004). A growing body of scholarship, catalyzed by early seminal studies (e.g., Adamson and Regan 1991), has inspired a complementary line of inquiry focusing on L2 speakers' capacity to acquire and manipulate variable TL patterns, generally referred to as Type II variation (see Howard et al. 2013, Bayley and Escalante 2022). The study of Type II variation is predicated on the core theoretical concept of the linguistic variable and the conviction that variation exhibits orderly or structured heterogeneity (Weinreich et al. 1968, Labov 1972), internalized by native speakers in the process of acquiring community-based norms during first language acquisition.

Because mastery of implicit constraints on variable use is claimed to be an integral component of native speaker sociolinguistic competence (Weinreich et al. 1968), it

follows that L2 speakers wishing to approximate native-speaker norms should acquire TL sociolinguistic patterns (Geeslin and Yim Long 2014). To what extent is this achievable? In addition to the much-debated effects of age on L2 acquisition (see e.g. Birdsong 2006 for a review), individual learner characteristics such as aptitude and motivation, together with an array of psychological and attitudinal considerations, are heavily implicated in the L2 learning process and its outcomes (Howard et al. 2013). A significant predictor of the L2 acquisition of TL vernacular norms involves sustained contact and interaction with TL speakers (see e.g. Sankoff et al. 1997, Blondeau et al. 2002, Nagy et al. 2003). In contrast with L2 classroom learners, who are reported to use informal variants of the TL less often than native speakers (Mougeon et al. 2004, 2010; Regan et al. 2009), high levels of contact between L2 acquirers and TL community members correlate with increased use of informal variants, often approaching native-speaker levels of frequency (Sankoff et al. 1997, Blondeau et al. 2002).

Yet even when L2 acquirers benefit from extensive naturalistic exposure to TL norms, the general consensus is that prolonged contact offers no guarantee that learners will faithfully reproduce the sociolinguistic conditioning governing variable use in the corresponding TL community. Rather, L2 learners are reported to exhibit incomplete mastery of the constraints on variation observed by TL speakers (Uritescu et al. 2004; see also Bayley and Regan 2004, Mougeon et al. 2004). Indeed, Schleef et al. (2011) go so far as to argue that the socio-cognitive complexity of replicating variable patterns in the TL is such that re-ordering or non-replication of constraints is a highly probable outcome for L2 speakers. Thus, in contrast with L1 acquisition, where children assiduously replicate the details of their variable input, L2 learners are not expected to "learn and reproduce linguistic forms, rules, and constraints with the accuracy and speed that children display" (Labov 2007: 349).

A key consideration among the factors believed to influence L2 acquisition of variable features of the TL is the nature of the variable being acquired, as well as the complexity of constraints conditioning variable use (Howard et al. 2013, Meyerhoff and Schleef 2014). Because each linguistic variable tends to be associated with a specific suite of constraints, it does not necessarily follow that learners who have mastered one ensemble of constraints will be equally successful in acquiring those associated with other variables (Mougeon et al. 2004). It has been suggested, for example, that variables lying at the syntax-discourse interface present acute learnability issues (see Sorace 2004). In this regard, Schleef (2017) observes that variable quotative expressions, involving the manipulation of both syntactic *and* pragmatic constraints, are less easily acquired by L2 speakers in the earlier stages of learning English than variables that are largely phonologically conditioned.

Yet another factor impinging on successful L2 learning concerns the relative stability of the TL variable being acquired. An understudied issue – and one which we confront in this article – concerns the ability of L2 learners to access and acquire ongoing linguistic change in the TL. This issue raises a number of key questions. How does L2 proficiency affect learners' ability to acquire change in the TL? And to what extent are L2 learners able, if at all, to emulate their TL peers in the incrementation of linguistic change, involving modifications in the "frequency, extent, scope or specificity" of a variable feature (Labov 2007: 346)?

Attempts to formulate empirical answers to these questions have encountered a number of obstacles. Chief among these is the extensive methodological infrastructure needed to conduct rigorous scientific investigations of L2 acquisition of ongoing changes in TL varieties. A first requirement is access to a substantial corpus of TL speech incorporating (minimally) an apparent-time dimension with which to identify and characterize the directionality of any change(s) in progress. A second is a sufficiently extensive corpus of L2 speech, collected contemporaneously in the same community representing the TL baseline variety, with which to assess the potential diffusion of change(s) from TL to L2 speakers.

Recent investigations of the L2 acquisition of the English quotative system (the locus of rapid, salient change (Buchstaller 2014), and the topic of our own case study below) illustrate the kinds of challenges that can arise as a result of differences in the learner populations studied, data limitations, and potential methodological confounds.

One unresolved issue concerns the extent to which findings abstracted from diverse learner populations can be marshalled to formulate robust generalizations about the nature of L2 acquisition. A cursory inspection of the recent literature on L2 acquisition of the English quotative system reveals that pertinent data have been drawn from disparate, and possibly incommensurable, learner groups, including recently arrived migrants (Meyerhoff and Schleef 2014, Corrigan 2020), speakers of English as a second language (ESL) and as a foreign language (EFL) (Davydova 2021), and learners in study abroad contexts (Magliacane and Howard 2019).

Paucity of data presents another major challenge for interpreting L2 patterns and venturing empirical generalizations. Targeting the acquisition of a range of English variables by young Polish migrants in the U.K., Meyerhoff and Schleef (2014) elicited just 43 instances of quotative expressions from their sample population. These included only 10 occurrences of the innovative *be like* variant (see also Choi and Diskin-Holdaway 2022 and Deng 2023 for similar data limitations), reported to be inaccessible to all but the most proficient L2 speakers. Additional evidence correlating the acquisition of *be like* with advanced L2 proficiency can be found in Diskin and Levey's (2019) investigation of the acquisition of quotative variation and change in Dublin English by Polish-born L2 migrants. Comparison of the detailed structure of variant choice in TL and L2 discourse revealed that although advanced learners had clearly acquired the *be like* quotative variant, they had not fully replicated the suite of constraints governing this variant in the corresponding TL baseline variety (Diskin and Levey 2019). Once again, potentially detracting from these key findings are the small speaker samples on which the investigation was based (Diskin and Levey 2019).

Even in studies based on more copious data, results may be difficult to interpret when they are drawn from L2 populations whose linguistic behaviour may not be easily contextualized in relation to relevant TL baseline varieties. Targeting the acquisition of the *be like* quotative by ESL and EFL university students in India and Germany respectively, Davydova (2021) reports that advanced learners of English were found to have not only acquired the innovative *be like* variant, albeit at lower rates than in age-matched TL populations, but, exceptionally, to have also largely reproduced the implicit set of constraints governing the selection of that variant. Davydova (2021) argues that in terms of frequency and constraint complexity, L2 mastery of the probabilistic constraints governing the use of *be like* is only moderately

taxing from a socio-cognitive perspective, in contrast with the acquisition of more intricately conditioned variables.

One potential drawback of Davydova's (2021) conclusions lies in the absence of any systematic comparisons of ESL and EFL speech varieties with the precise TL sources from which L2 learners presumably acquired patterns of quotative variation and change. Because *community-specific* speech data produced by the original agents of transmission is not factored into Davydova's (2021) analysis, it is not possible to determine which varieties, exactly, the L2 speakers in question were exposed to, let alone clarify the precise mechanisms that facilitated the diffusion of TL variable patterns to L2 speakers. These methodological infelicities serve as an important reminder that, since *be like* is situated within a larger variable system, where there are amply documented constraints on its use (e.g., Tagliamonte and Hudson 1999; Tagliamonte and D'Arcy 2004, 2007; Buchstaller and D'Arcy 2009; Tagliamonte et al. 2016), some amount of sustained face-to-face interaction with TL adopters of this innovation is required to enable it to diffuse across populations (see e.g. Dion and Poplack 2007).[1]

In the following sections, we detail how a community-based perspective, judiciously integrated with a comparative sociolinguistic approach, can be deployed to transcend some of the methodological problems that have hampered previous investigations of L2 acquisition of variation and change in TL varieties.

3. Choice of community

The geographical focus of our research project is the Canadian National Capital Region. This includes the cities of Ottawa (Ontario) and Gatineau (Quebec), as well as their surrounding municipalities, essentially constituting a single urban conglomeration.

This area is identified as a "natural laboratory for language contact" (Poplack 1989: 413) because it is situated on the Ontario-Quebec provincial border, where English is the majority language in Ontario whereas French is the official and majority language in Quebec. According to recent census information (Statistics Canada 2021), 58.2% of the population in the city of Ottawa claim English as a mother tongue, whereas only 12.5% of the population are French mother-tongue claimants. Knowledge of both official languages is reported by 36.4% of the local population. On the other side of the provincial border, across the Ottawa River, the sociolinguistic context presents a sharp contrast: 71.1% of the population in the city of Gatineau claim French as a mother tongue, whereas only 11.8% of the local population are English mother-tongue claimants. Another striking difference is that knowledge of both official languages is much higher than in Ottawa, reported by 65% of the local population.

Prolonged and intense contact between the two official languages has fuelled the popular belief that many of the putatively distinctive features of Canadian French are the result of English influence, the numerically and politically dominant language

---

[1]The notion that the media play a role in the long-distance diffusion of *be like* (e.g., Trudgill 2014) finds scant support in empirical research to date (see Dion and Poplack 2007).

(Poplack et al. 2013). In the public imagination, the perceived anglicization of Canadian French is routinely construed as a harbinger of the eventual cultural and linguistic assimilation of Canadian francophones by the anglophone majority (Poplack et al. 2019). Consequently, in areas where French is spoken as a minority language, the L2 acquisition of (advanced) proficiency in English by francophones is invested with considerable ideological freight due to the spectre of language shift that it inevitably provokes. We return to these issues below, as they have a potential bearing on affective orientations towards, and motivations for, acquiring English as a second language.

4. Choice of speakers

In constituting a corpus of spontaneous L2 speech, our sampling protocol targeted francophone adults who had acquired Canadian French as their primary language from native francophone parents/caretakers and who had been raised in the Canadian National Capital Region. In addition to native ability in French, another precondition for inclusion in the sample required speakers to have completed their mandatory schooling in French-speaking educational establishments. Speakers who had been raised bilingually (i.e., French and English) from birth were excluded from the study.

Between 2018 and early 2022, we recorded a total of 29 speakers meeting the basic sampling requirements detailed above.[2] The distribution of sample members by speaker sex and age is shown in Table 1 below.

As is evident from Table 1, the majority of the speakers are relatively young, enabling us to test whether speakers within the 19–33 age range would be likely to acquire any ongoing change in the TL from their anglophone peers. The number of older speakers in the dataset is small, and restricted to female speakers only. We nonetheless capitalize on the limited apparent-time dimension incorporated into the sample to explore the potential effects of age on the L2 acquisition of TL patterns of variation and change.

We additionally recorded speakers' occupational status, although in the relatively small sample at our disposal, it was not feasible to utilize speaker occupation as a stratifying measure. At the time of data collection, many of the younger speakers were engaged in post-secondary education, although several were also employed in the hospitality services, skilled trades and in retail, as well as in agriculture and technical positions. Among the older speakers, over half were retired, whereas others were either engaged in skilled employment or were on extended leave from work.

Several parameters relevant to L2 acquisition were surveyed using language background questionnaires. These allowed us to probe exposure to formal instruction in English, the language (French, English) used most often in an individual's daily life, and frequency of English-language use at home, at work, and in speakers' local neighbourhood, as well as for purposes of socializing.

---

[2]The bulk of the data was recorded between 2018 and 2020. Owing to the COVID-19 global pandemic, data collection was temporarily interrupted and four speakers were recorded somewhat later, in early 2022. Our inability to record older (50+) males meeting the sampling requirements necessarily limits the analysis of quotative variation and change among older L2 speakers and constrains our interpretation of age-related differences within the L2 cohort.

Table 1. Distribution of L2 sample members by sex and age

	19–33 years	50–77 years	Total
Females	10	7	17
Males	12	0	12
Total	22	7	29

Most speakers reported having been initially exposed to formal instruction in English in either grades three or four of their compulsory schooling (i.e., between the ages of 8 and 10). Content analysis of the L2 interview data additionally confirmed that sporadic exposure to English in L2 speakers' neighbourhood and via the broadcast media (e.g., television) had been a common childhood experience.

Only three speakers reported that they used English more often than French in their daily lives. The primary domain of regular English-language use was in speakers' work/study environments, with English being least frequently employed at home, where French predominated. As we would expect, speakers (N = 9) residing on the Quebec side of the provincial border reported generally very low levels of English language exposure in their neighbourhood, whereas those (N = 20) residing in Ontario reported much higher rates of exposure to English.

Even more revealing, as shown in Table 2 below, is the fact that only two speakers in the L2 sample reported having no anglophones in their individual friendship networks. Fifty-five percent of L2 speakers claimed that native anglophones comprised 50% or more of their personal social networks, pointing to many L2 speakers' close affiliation with, if not integration into, the local TL community.

Interviews conducted in English by native anglophones with L2 speakers averaged 55 minutes in length. With no individual recording lasting less than 40 minutes, all speakers were able to sustain an interview in their second language, with varying degrees of fluency.

Notoriously difficult to operationalize, but of primordial importance for understanding L2 acquisition and its outcomes, is the controversial notion of L2 proficiency. Recognizing that even relatively advanced L2 speaker groups subsume a

Table 2. Distributional analysis of L2 speakers' self-reports of the estimated proportion of native anglophones in their personal social networks

Estimated proportion of anglophones in L2 speakers' social networks	No. of speakers in L2 sample	Percentage of speakers in L2 sample
0%	2	7
1%-25%	3	10
25%-50%	8	28
50%-75%	13	45
75%-100%	3	10
Total	29	

range of individual proficiency levels, we initially sought to determine proficiency via speaker self-ratings of their English production and comprehension skills on a scale of 1 to 10. Table 3 summarizes speakers' self-assessments of English proficiency according to their ability to produce and comprehend the spoken language.

As a first general observation, it is apparent that, in the aggregate, L2 speakers rate their comprehension skills more favourably than their production skills, in keeping with the general assumption that in acquiring language, comprehension precedes production (see e.g. Tasseva-Kurktchieva 2015). Still, only five speakers (17% of the sample) rate their production skills as intermediate, with the majority qualifying their oral proficiency in English as either "very good" or "excellent".

We caution, however, that self-reported proficiency is a somewhat imprecise measure (Meyerhoff and Schleef 2014) because the capacity to self-assess language use is unevenly distributed among the general population (Romaine 1995). As such, self-reported proficiency is prone to under- or over-estimation (Poplack 2018). To mitigate this problem, we calculated for each speaker a composite measure of English proficiency, the *Cumulative English Proficiency Index* (CEPI), inspired by procedures described in earlier research on language contact (see e.g. Poplack et al. 2006, Poplack 2018, Torres Cacoullos and Travis 2018). Following Torres Cacoullos and Travis (2018), individual CEPI scores were calculated on the basis of data abstracted from speaker self-assessments of English language proficiency, information gleaned from individual language background questionnaires administered to all L2 speakers (e.g., relative use of English at home, work, in the local neighbourhood and for purposes of socializing, as gauged from scalar responses), and content analysis of speakers' actual L2 production data. This composite approach generated individual CEPI scores, calculated from the various complementary measures we operationalized, ranging from a low of 0.450 to a high of 0.863. The resultant spectrum of scores enabled us to subdivide the L2 speaker sample into four proficiency bands, as shown in Table 4 below.

We make no claim that CEPI scores are holistic, absolute measures of English language proficiency. Rather, scores are intended to be interpreted relative to one another, and offer a general assessment of the vitality of English in each L2 speaker's communicative repertoire (see Poplack 2018). We draw on these scores in our case

Table 3. L2 speakers' self-reported proficiency ratings according to English-language production and comprehension skills

Self-reported rating	English production skills		English comprehension skills	
	No. of speakers	% of sample	No. of speakers	% of sample
Intermediate	5	17	0	0
Very good	16	55	12	41
Excellent	8	28	17	59
Total	29		29	

Table 4. Cumulative English Proficiency Index (score ranges shown in parentheses)

	No. of speakers	% of sample
Low (0.450–0.494)	3	10
Mid-low (0.575–0.694)	9	31
Mid-high (0.700–0.788)	11	38
High (0.800–0.863)	6	21
Total	29	

study below to explore their correlation with L2 speakers' ability to acquire variation and change in the TL.

5. Attitudes towards the acquisition of English as a second language

A growing body of scholarship indicates that individual attitudes towards languages in bilingual settings, as well as the relative social prestige with which contact varieties are imbued, can influence motivations for L2 acquisition and levels of L2 attainment (see e.g. Gardner and Lambert 1959, 1972; Dörnyei et al. 2006).

Researchers have operationalized a broad distinction between instrumental and integrative orientations to second language learning. An instrumental orientation to second language learning is one which is reportedly based on the "utilitarian value" of L2 acquisition (Gardner and Lambert 1959: 267), characterized by the social and economic gains that knowledge of a second language may facilitate. By contrast, an integrative orientation appears to be motivated by the desire to learn more about the TL group (Gardner and Lambert 1959) and to develop affiliative ties with that group (Baker 1992).

In the Canadian context examined here, where, as noted earlier, language choice is a socio-politically fraught issue (see e.g. Taylor et al. 1977, Bourhis 1984), we might conjecture that a broadly instrumental orientation to the acquisition of English (e.g., as an occupational necessity) would prevail in the targeted L2 population. Alternatively, given the proportion of francophones in our L2 sample with large numbers of native anglophones in their personal social networks (see Table 2), it is not unreasonable to expect that an integrative orientation to the acquisition of English would be discernible among speakers who have close personal ties with TL community members.

Accessing community-based language attitudes presents a number of difficulties. A major challenge hinges on the conversion of the subjective notion of attitude into something which is objectively measurable (Romaine 1995). Among the problems that arise when individuals are explicitly asked about their attitudes towards language varieties (and their speakers) is that direct elicitation tends to evoke socially desirable responses which may not be fully congruent with individuals' genuine feelings towards a particular variety or speaker group (Baker 1992, Kircher 2009).

To attenuate some of the difficulties with accessing speaker attitudes, we exhaustively extracted from the L2 interview data (circa 26 hours of audio recordings) all
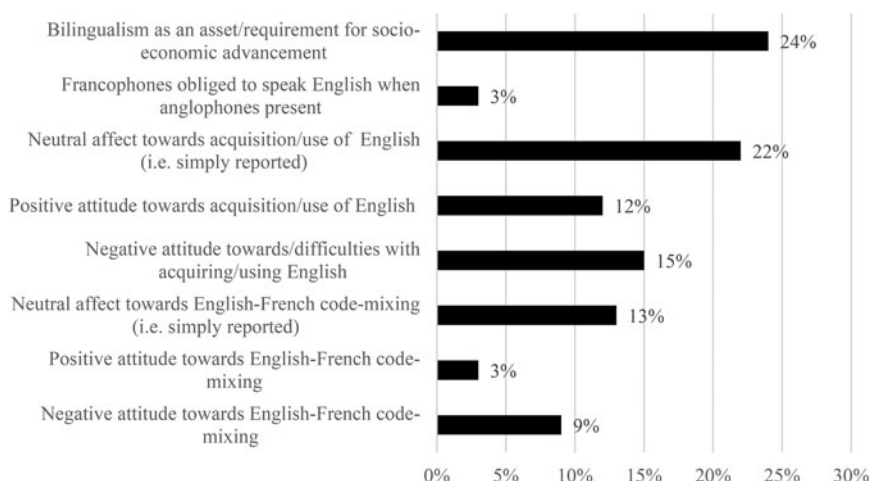
Bilingualism as an asset/requirement for socio-economic advancement — 24%

Francophones obliged to speak English when anglophones present — 3%

Neutral affect towards acquisition/use of English (i.e. simply reported) — 22%

Positive attitude towards acquisition/use of English — 12%

Negative attitude towards/difficulties with acquiring/using English — 15%

Neutral affect towards English-French code-mixing (i.e. simply reported) — 13%

Positive attitude towards English-French code-mixing — 3%

Negative attitude towards English-French code-mixing — 9%

Figure 1. Proportional analysis of attitudes (Total N = 260) towards English acquisition and use as well as French-English code-mixing and bilingualism

remarks relating to English acquisition and use, as well as those pertaining to contact between French and English. Our extraction protocol targeted comments that were negative, positive or that had no discernible affective orientation. Extracted comments were subsequently grouped according to several broad themes reflecting attitudes towards facets of English acquisition and use, as well as assessments of French-English code-mixing and bilingualism. The resultant taxonomy, though based on broad-brush categories, has the capacity to illuminate major preoccupations within the targeted L2 community (see also Poplack 2018).

As shown in Figure 1, almost a quarter of the responses characterize knowledge of both official languages as a front-end investment for securing better employment opportunities, as exemplified in (1):

(1) …because there's an advantage to being bilingual. Uh my kids would not have the jobs they had or that they will have because […] if they weren't bilingual. And I truly believe that uh you will be hired being bilingual, you will be hired first… (L2/010/15510)[3]

References to the acquisition and use of English are predominantly expressed with neutral affect. Positive attitudes towards the acquisition and use of English, illustrated in (2), are slightly less prevalent than those expressing negative affect. Negative evaluations tend to focus on lack of confidence in English-language production skills, as

---

[3]Examples are reproduced verbatim from the spoken corpora. Codes in parentheses refer to the corpus from which the example is drawn (L2 = Second Language Corpus of English; TL = Target Language Corpus of English; FL1 = French First Language Corpus); the unique speaker identifier; and the co-ordinates (line number or time stamp) of the utterance in the respective corpus.

in (3), as well as occasional resentment towards having to use the language, illustrated in (4).

(2) that's how it was for me, like you feel cooler when you're speak— speaking English than when you— than when you do when you're speaking French. (L2/008/13342)

(3) I don't think my English is that bad, but if you had met me like four years ago, or maybe five, six years ago, I was barely able to maintain a basic conversation in English. (L2/002/4879)

(4) …but uh, before it used to really bother me, I hated speaking English, and I hated everything about— everything that was English, like I actually had a phase in my life that like anything English I'd like cut it off. (L2/015/1711)

Although representing only a very small proportion (3%) of expressed attitudes, criticism is also levelled at the need to accommodate generally lower levels of bilingualism among anglophones by having to switch to English in contexts where speakers of both languages interact, but anglophones are profoundly in the minority:

(5) And I hate the double standard where in th— when there is one English person in the room and everybody else is French, everybody has to start speaking English. (L2/004/6658)

Given the well-documented social, cultural and institutional antipathies towards the supposedly insidious effects of English on the structural integrity of Canadian French (Papen 2022), it is notable that metalinguistic observations on English-French language contact are mostly neutral. Still, only a relatively small proportion (3%) of comments assess the consequences of language contact in a positive light, with a larger proportion (9%) tending to denigrate English influence on spoken French, consonant with dominant societal ideologies:[4]

(6) …but my brother yeah, like his friends are super anglophones and like yeah so… He uses a lot of English words, more than I do. And like his French— his French is way worse than mine (laughter) yeah (L2/017/3291)

Summarizing, although we caution that verbally expressed language attitudes may not always coincide with speakers' deeply rooted feelings towards varieties and their associated communities (Giles et al. 1977), the data in Figure 1 suggest that knowledge of English, in addition to French, is generally valued for its instrumental capacity to foster social and economic advancement. This attitudinal orientation is perhaps unsurprising given the current status of English as a global lingua franca and its association with upward social mobility (Kircher 2009). It also accords with much official rhetoric in Canada extolling the value of bilingualism as an asset in a competitive

---

[4]Such attitudes are by no means predictive of speaker behaviour. Speakers who explicitly denigrate English lexical influence on French may actively engage in this very behaviour themselves. Moreover, they may openly admit to doing so.

labour market. It is nonetheless clear that certain francophones in our sample exhibit integrative motivations for acquiring and using English. Speakers alluded, for example, to its socio-symbolic utility for indexing allegiance to adolescent youth cultures and strategically opposing the language of the school environment, pitting English as the language of "cool" against French, the language intimately associated with institutional status, adult authority figures and formal learning:

(7) Uh I think if you're starting to speak French you're associated with school and associated with the profs, the old people that speak French, and all you know all of what's considered cool is we're… we're grown up on uh social media and these TV shows which is all mostly English and I think that's what was associated with being cool. (L2/008/13335)

Occasional sentiments of linguistic insecurity expressed by speakers in relation to their native French raise the possibility that francophones with advanced English-language proficiency might shift largely or entirely to English rather than living with a stigmatized linguistic identity (Kircher 2009). We emphasize here, however, that we detect no indication in speakers' expressed attitudes of any willingness to abandon French and embrace English.[5] On the contrary, as indicated in (8) below, speakers ascribe positive affective value to their native French, highlighting it as an intrinsic symbol of their own cultural and personal identity.

(8) Cos like for me it's French cos it's like I… I really like uh identify myself like in French you know like but um for anyone like any culture that would be lost or start to be assimilated like that for me like it's a big loss for humanity. (L2/002/4411)

6. Corpus compilation

6.1 L2 data collection and computerization

Community-based L2 speech data were collected by trained research assistants by means of sociolinguistic interviews, designed to promote the use of vernacular speech reported to yield "the most systematic data for linguistic analysis" (Labov 1984: 29).

In line with previous research utilizing corpus compilation as a conduit to detailed quantitative analysis, digital recordings of L2 production data were fully transcribed and computerized. As the creation of any sizable corpus of transcribed speech necessitates "a major trade-off between the size of the data base and level of detail of the transcription" (Poplack 1989: 430), we used standard orthography to represent variation resulting from phonetic or phonological processes, regardless of how forms were actually produced. We reason that any empirical study wishing to target phonetic or phonological variables in the data would rely primarily on the original digital audio recordings, rather than the transcriptions, as the principal source of

---

[5]Further evidence emerging from additional analysis (not shown here) of L2 speaker attitudes towards their native French revealed that 69% (N = 177) of attitudes expressed either a neutral or positive stance towards local varieties of French, including the desire to see those varieties maintained.

information (see Poplack 1989). By contrast, any processes affecting the omission (or addition) of an entire morpheme (e.g., the variable omission of plural -*s*) were faithfully respected in the resultant transcriptions, as were speaker reformulations, hesitations, disfluencies, syntactic anomalies, other-language insertions and code-switches.

The allocation of a unique numeric identifier to speaker utterances enabled us to link individual speech production data with a relational database containing detailed demographic information (e.g., speaker sex, age, occupation, contextual information relating to L2 acquisition of English, preferred language use in different domains, etc.) necessary for the investigation and interpretation of sociolinguistic patterns in L2 speech (see Labov 1984).

Following the implementation of correction protocols to the data, some 277,000 words were retained in the final transcribed version of the corpus (see Table 5 below). This corpus, together with its associated demographic database, constitutes a fully automated resource that can be used to locate diverse lexical, discursive and morphosyntactic phenomena with maximum efficiency, lending itself to the study of a range of research questions impinging on L2 acquisition.

6.2 Additional corpora

The *Ottawa English Corpus* (OEC) compiled in 2019, comprising recordings of 34 native anglophones born and raised in the Ottawa-Gatineau metropolitan area, serves as the TL community baseline against which L2 usage patterns are calibrated (see Table 5 below).[6] The data collection period of the TL benchmark variety coincides with the time frame in which the L2 data were gathered, enhancing the descriptive adequacy of the comparisons we draw between TL and L2 speaker groups. Stratified by speaker sex and age (19- to 28-year-olds versus 40- to 72-year-olds), the TL corpus enables us to situate ongoing change in an apparent-time perspective.

A third corpus based on recordings of vernacular French elicited from 20 francophones who also contributed to the L2 English corpus allows us to compare the conditioning of variable linguistic structures in speakers' L1 and L2. Stratified by age and speaker sex, this corpus includes some 228,000 words of transcribed speech. The primary motivation for constructing this resource derives from the longstanding recognition that speakers' L2 is potentially susceptible to influence from their native language (see e.g., Weinreich 1968). Typological similarities shared by a speaker's native language and the TL are claimed to be especially hospitable to cross-linguistic influence (Backus 2005, Thomason 2013), even in cases where no overt other-language material is actually transferred (Odlin 2003). Previous research hypothesizes that L2 speakers will favour TL variants that have a morphological and semantic equivalent in their native language (Mougeon et al. 2010). If such inter-linguistic influence is present in our data, our research design should enable us to detect it.

Table 5 below provides a summary of the corpora incorporated into our research design for comparative purposes. These corpora are a rich repository of narratives of personal experience, discussions of language and personal opinions, as well as other lengthy extracts of casual speech reflecting spontaneous everyday usage.

---

[6]This corpus is part of a much larger monitor corpus, the compilation of which was originally initiated in 2008.

Table 5. Characteristics of the corpora used to investigate the L2 acquisition of English by Canadian francophones (Notes: OEC = *Ottawa English Corpus*; CNCR = Canadian National Capital Region)

	L2 Corpus of English	TL Corpus of English (OEC)	L1 Corpus of French
N. hours of recordings	26 hours	26 hours	17 hours
N. transcribed words	277,586	Not transcribed	228,530
N. speakers	29	34	20
Speaker age range	19–77 years	19–72 years	19–77 years
Data collection	2018–2022 (CNCR)	2019 (CNCR)	2018–2020 (CNCR)

In the following sections, we illustrate how we can capitalize on the wealth of vernacular data at our disposal to elucidate the L2 acquisition of TL patterns of variation and change.

7. Case study: Acquiring variation and change in the target-language community
7.1 Context

The variable expression of quotation is an exemplary candidate for investigating the L2 acquisition of community-based TL patterns of variation and change, since both the TL, English, and L2 speakers' L1, French, feature a set of partially overlapping forms used "to introduce reported speech, sounds […] and thought by self or other" (Buchstaller 2006: 5).[7]

The quotative systems of Canadian English and French have both witnessed the rise and increasing productivity of the non-canonical variants *be like* and *être comme,* exemplified in (9) and (10) respectively. These innovations alternate with more established variants, as illustrated below in (11)–(14).

(9) And so my dad *was like,* "Oh, you— you should move here da da da" (TL/002/11:33)

(10) J'*étais comme*, "Calme-toi, je suis en première année" (L1/005/11581)
       'I was like, "Calm down, I am in first year" '

(11) And I *go*, "Yeah, but it makes it look like you love her more" (TL/003/21:07)

(12) And so my parents *said*, "Hey Electra, if you wanna do something, you can do one too" (TL/005/09:12)

(13) Pis j'*ai fait*, "Ah ben! Je vais retourner chez nous" (L1/002/4154)
       'And I went, "Ah well! I'm going back home" '

(14) ils *ont dit*, "Ah, on peut pas faire d'exceptions parce que sinon il y aurait trop d'exceptions" (L1/023/154)
       'They said, "Ah, we can't make exceptions because otherwise there would be too many exceptions" '

---

[7]In a small number of cases, a quotative expression can introduce a performed gesture, but these are omitted from consideration here because such instances cannot be reliably reconstructed and interpreted from the digital audio files alone.

Since its earliest recorded attestations in the 1980s (Butters 1982), the innovative *be like* quotative variant has diffused rapidly throughout the anglophone world, to the extent that it is now the unrivalled exponent of the quotative system used by English-speaking youth (see e.g. Tagliamonte and D'Arcy 2007, Buchstaller and D'Arcy 2009, Durham et al. 2012, Tagliamonte et al. 2016, Hunt Gardner et al. 2021).

Though less extensively investigated (see e.g. Levey et al. 2013*)*, the emergence of *être comme* in Canadian French appears to post-date the vertiginous rise of its English counterpart. What little quantitative research there is on *être comme* in Canadian French traces early occurrences to adolescent speech in the mid-1990s (Vincent 2005). Recent investigations (e.g., Levey et al. 2013) drawing on real- and apparent-time evidence offer compelling indications that, like its English equivalent, the increase in the use of *être comme* in urban varieties of Canadian French qualifies as another possible instance of rapid language change, unfolding over the past three decades.

The parallel historical origins of *be like* and *être comme*, their overlapping trajectories of change, as well as similarities in surface form and contexts of use, conspire in suggesting that *être comme,* as a seemingly more recent development than *be like*, qualifies as a possible case of pattern replication (see Matras and Sakel 2007) or replica grammaticalization (see Heine and Kuteva 2005) triggered by the sustained contact of French with English. Further support in favour of this scenario can be found in Sankoff et al.'s (1997) study of Anglo-Montreal French (i.e., L2 French). This investigation revealed that *être comme* was well attested in L2 speech but was absent, as far as the authors could determine, from the corresponding local French TL baseline variety, suggesting that its presence in L2 French was the product of transfer effects traceable to L2 speakers' native English.

But the emergence of *être comme* in L1 varieties of French can be attributed equally well to other causal mechanisms. There is ample cross-linguistic evidence of the development of quotative expressions from source constructions encoding similarity and manner (see e.g. Güldemann 2008). The existence of such a pathway suggests that the origins of *être comme* could reside in internal as opposed to externally motivated change. Additional support for this possibility can be adduced from multiple reports of developments paralleling the emergence of quotative *be like* in a range of typologically related and unrelated languages (see Buchstaller and Van Alphen 2012, Buchstaller 2014). A third possibility that cannot be ruled out is that *être comme* is the product of multiple causation (Thomason 2001), with sustained contact between English and French possibly triggering internal change in the latter (see Heine and Kuteva 2005), although straightforward tests of this hypothesis are notoriously difficult to operationalize.

However the genesis of *être comme* in Canadian French is ultimately explained, its formal and functional parallels with English *be like* instantiate precisely the kinds of preconditions believed to be hospitable to cross-linguistic transfer effects. In the following section, we consider how a variationist framework can engage with this issue by illuminating the structural context hosting quotative variation in each of the datasets depicted in Table 5.

7.2 A variationist perspective

A key working principle informing the sociolinguistic investigation of natural language data is that everyday speech is inherently variable and rule-governed (Weinreich et al.

1968). Of particular importance is the recognition that linguistic variability is conditioned by multiple social and linguistic factors. These factors are probabilistically associated with the distribution of competing variants in discourse. Quantitative modelling of these probabilistic associations affords insights into the detailed structural configuration of the variable system hosting competing variants (see Poplack and Tagliamonte 2001). The resulting structural information can then be compared across speaker cohorts and varieties, as we do below, with the aim of pinpointing affinities and differences between them. Conflict sites or areas of functional, structural and/or quantitative differences between comparison varieties (Poplack and Meechan 1998) assume particular importance in adjudicating between competing hypotheses about the source and nature of the processes constraining observed patterns of variation (Poplack and Levey 2010).

We muster this structural information to test our working hypothesis that the youngest and most proficient L2 speakers (as assessed from their CEPI scores) will be frontline users of the innovative *be like* quotative and will mostly closely approximate TL constraints on variant selection. We also examine whether L2 speakers' use of *être comme* in their native French has any appreciable effect on the use *be like* in their L2, as inferred from comparisons of the ordered set of constraints governing the selection of *être comme* in their L1 and *be like* in their L2.

7.3 Method

In accordance with the principle of accountable reporting (Labov 1972), every expression introducing direct quotations reporting speech, thought or sound effects (Buchstaller 2006) was extracted from each dataset summarized in Table 5. Quotative expressions reporting indirect speech and excerpts of written language were excluded from the analysis, as were incomplete and anomalous utterances.

To optimize the comparison of our results with those of previous studies of quotative variation and change, we focus on the standard extra-linguistic predictors of age and sex as well as the following linguistic constraints governing variant selection: grammatical person, content of the quote, mimesis, and temporal reference.

With regard to grammatical person, we distinguish first-person (15) from third-person contexts (16), which, in turn, are differentiated from all remaining environments (e.g., second person and null subjects), in order to test the previously reported association of *be like* (and *être comme*) with first person referents (see e.g. Buchstaller and D'Arcy 2009, Levey et al. 2013):

(15) so I *was just like*, "Whatever I'll go back to normal" (L2/022/1875)
(16) She's *like*, "Well I want a refund. I want a refund" (L2/018/5822)

Another well-attested constraint on variant choice concerns the content of the quote. Here we explore the distinction, foregrounded in previous research, between reported speech, as in (17), and internal dialogue/inner states, in (18) (see Ferrara and Bell 1995). The preferred association of *be like* with internal dialogue/inner states as opposed to reported speech is a robust finding documented in earlier studies (also reproduced in the case of *être comme*; Levey et al. 2013), although weakening or reversal of that constraint has occasionally been noted (see e.g. Tagliamonte and D'Arcy 2004, 2007).

(17) my friend *was like*, "Oh if you do that you're gonna be working with a computer" (L2/023/9792)

(18) Um so I started thinking it was weird that I would always get caught so I *was like*, "I'm just not gonna do bad things anymore" (L2/009/11620)

Less extensively researched is the effect of mimesis (Güldemann 2008) on variant choice. Mimesis is understood here to refer to the use of sound symbolism and the modulation of supra-segmental features (e.g., loudness, stress and pitch) in the quoted extract to construct stylized or dramatized renditions of human dialogue. We contrast quoted extracts containing mimetic re-enactments of dramatized dialogue, as in (19), with those showing no evidence of mimetic effects, as in (20). Mimesis-encoding contexts are reported to be particularly hospitable to new quotative expressions such as *be like* and *être comme* (Levey et al. 2013, Buchstaller 2014).

(19) So my parents *are like*, "Hell yeah, are we selling our house?" (L2/018/6012)

(20) And I would hold the phone and I'*m like*, "Do you want me to call mom?" (L2/006/9571)

The contribution of temporal reference to variant choice is reported to exhibit locally idiosyncratic effects across varieties of English (Buchstaller and D'Arcy 2009), although earlier studies (Levey et al. 2013) targeting English and French in the Canadian National Capital Region indicated that *be like* and *être comme* both showed a preference for the conversational historical present (CHP) in the discourse of younger speakers. To detect whether similar correlations obtain in our data, we distinguish quotatives encoded with non-past-tense morphology with past temporal reference (i.e., the CHP) as in (21); quotatives encoded with non-past-tense morphology with present temporal reference, as in (22); and quotatives encoded with past tense morphology with past temporal reference in (23). As the temporal reference system of French is not identical to that of English, our coding protocol also accounted for tense/aspect forms relevant to the expression of quotative variation in French, such as the imperfect, exemplified in (24). All residual tense/aspect distinctions in both languages (e.g., the past perfect and future, etc.) were grouped under 'other'.

(21) My friend gave me his jacket and he'*s like*, "Ah I can't take this man" (L2/004/6610)

(22) I think most people *are just like*, "Okay so we're there, so we're there" (L2/008/13263)

(23) I *was like*, "Are you serious?" (L2/022/1968)

(24) Mes parents *étaient comme*, "Non, tu t'en vas pas là tout seule" (L1/005/11885) 'My parents were like, "No, you are not going there by yourself" '

7.4 Results

7.4.1 Variant distributions

Table 6 gives a breakdown of variant distributions by speaker sex and age in the TL benchmark corpus. Occurring at an overall rate of 64%, *be like* is by far the dominant

Table 6. Distribution of quotative variants by speaker sex and age in the Target Language Corpus of English (OEC)[8]

| | 19–28 | | | | 40+ | | | |
| | Females | | Males | | Females | | Males | |
Variant	N	%	N	%	N	%	N	%
Be like	389	72	197	53	54	18	32	20
Say	39	7	95	26	142	46.5	70	44
Discourse *like*	25	5	17	5	5	2	9	6
It's like	24	4	7	2	5	2	2	1
Zero	22	4	13	3	17	5.5	20	12
Go	2	0	5	1	12	4	4	2
Think	2	0	2	1	29	10	5	3
Other	36	7	36	10	41	13	18	11
Total	539		372		305		160	

variant in the discourse of younger (19- to 28-year-old) anglophone speakers. Only in the speech of younger males does any other single variant (i.e., *say*) account for more than 10% of the variable context. Closer inspection of the younger age cohort additionally reveals an asymmetrical pattern of sex differentiation, with females' use of *be like* notably exceeding that of males (see also Ferrara and Bell 1995, Tagliamonte and D'Arcy 2004, 2007). Inspection of the data from the older (40+) age group indicates that *be like*, once the preserve of the under-40s (see Ferrara and Bell 1995, Tagliamonte and D'Arcy 2007), is percolating up the age spectrum, pointing to the participation of older (40+) speakers in the proliferation of this innovation.[9] Summarizing, the data shown in Table 6 indicate that *be like* is well en route to colonizing the quotative system of 19- to 28-year-olds and has made noticeable inroads into the discourse of older speakers, consistent with its characterization as a rapid language change phenomenon (Buchstaller 2014).

To what extent have L2 speakers of English acquired this change in progress? Table 7 compares variant forms and distributions by speaker sex and age in the L2 dataset.

---

[8]In Tables 6–9, the low-level variants, discourse *like* and *it's like,* and their French equivalents, discourse *comme* and *c'est comme*, are tabulated separately from *be like* and *être comme* respectively and are not amalgamated with tokens of *be like* or *être comme* in any of the statistical analyses associated with Figure 2 and Tables 10–13. There is evidence to suggest that *it's like* and *c'est comme* have distinct functional profiles and are used at very elevated rates to report internal dialogue (see Tagliamonte and D'Arcy 2004, Levey et al. 2013).

[9]In Table 6, the sex effect for 40+ speakers is spurious: for males, *be like* is concentrated in the 40- to 50-year-old age range, whereas for females, it is additionally found, albeit less frequently, in the speech of 50- to 60-year-olds. These results contrast with those reported in Levey et al. (2013) based on data collected in the same location between 2008 and 2010: out of more than 240 quotatives produced by speakers aged 40+, not a single instance of *be like* was recorded.

Table 7. Distribution of quotative variants by speaker sex and age in the L2 Corpus of English

| | 19–33 | | | | 50–77 | |
| | Females | | Males | | Females | |
Variant	N	%	N	%	N	%
Be like	244	75	103	48	16	7
Say	9	3	23	11	133	57
Discourse *like*	24	7	18	8	5	2
It's like	19	6	11	5	3	1
Zero	10	3	18	8	28	12
Go	1	0	7	3	14	6
Think	1	0	2	1	10	4
Other	19	6	32	15	24	10
Total	327		214		233	

We first observe similar variant inventories in L2 and TL discourse. A second important finding is that the overall rate of *be like,* accounting for 64% of quotative use in the L2 speech of 19- to 33-year-olds, is identical to that found in the TL baseline variety. A similar pattern of sex differentiation also obtains in the younger L2 cohort. In the L2 data, we also note that female speakers use *be like* at a rate that marginally surpasses its frequency in the discourse of females in the TL benchmark variety. Although older (50- to 77-year-old) female speakers in the L2 corpus make restricted use of *be like*, its limited occurrence in their speech suggests that the grammar of older L2 speakers is not entirely immune to this innovation.

These results would appear to offer provisional evidence indicating that younger L2 speakers have successfully acquired ongoing change from the local TL community. But to what extent is this accomplishment mediated by L2 proficiency? Focusing on the younger 19- to 33-year-old speakers, the primary L2 adopters of *be like*, Table 8 compares variant inventories and distributions in the L2 of low/mid-low proficiency speakers (CEPI score range = 0.450-0.694) and mid-high/high proficiency speakers (CEPI score range = 0.700-0.863).

Be like is the primary exponent of the quotative system in both proficiency groups. But while it virtually saturates the quotative system of the mid-high/high proficiency speakers, effectively relegating all other variants to a decidedly peripheral role in the system, it accounts for less than half of the quotatives used by low/mid-low proficiency speakers, who make comparatively greater use of alternative options such as *say,* zero as well as assorted infrequent constructions grouped under *other*.

If, as the current literature suggests, the acquisition of, and participation in, ongoing change in the TL is an exacting task for L2 speakers (see e.g. Bayley and Preston 2008, Meyerhoff and Schleef 2014), it stands to reason that degree of proficiency in English (as gauged from CEPI scores) will be a major determinant of L2 speakers' ability to meet that challenge. Furthermore, as Sankoff et al. (1997) point

Table 8. Distribution of quotative variants according to L2 proficiency

	Low/Mid-Low Proficiency (CEPI 0.450–0.694)		Mid-High/High Proficiency (CEPI 0.700–0.863)	
Variant	N	%	N	%
Be like	73	46.5	274	72.5
Say	18	11.5	14	4
Discourse *like*	12	8	30	8
It's like	7	4.5	23	6
Zero	14	9	14	4
Go	5	3	3	1
Think	2	1	1	0.3
Other	26	17	25	7
Total	157		384	

out, the capacity to acquire and use vernacular features of the TL that are not explicitly transmitted via formal instruction is a hallmark of speaker fluency as well as an indication of integration into the local TL speech community.

To what extent may structural similarities between L2 learners' native language, French, and the TL facilitate the L2 acquisition of quotative variation and change? To address this question, we first turn to the distribution of quotative variants in L2 speakers' native French, as shown in Table 9. With the exception of the low-level variants *faire* 'to do' and *se dire* 'tell oneself/think', all remaining options listed in Table 9 have equivalents in the corresponding TL data, and exhibit similar, though not identical, rates of occurrence. Of particular interest is the fact that although *être comme* clearly dominates the quotative system used by younger francophones, accounting for 51% of the variable context, it occurs somewhat less often than *be like* in the corresponding L2 and TL datasets. Just as in the case of *be like*, we observe a female lead in the use of *être comme*, as well as noting that it too has made very limited inroads into the speech of older (50+) females.

These distributional results lend weight to claims that L2 speakers will have a predilection for TL variants which have formal and distributionally robust counterparts in their first language. In order to verify whether there is any significant correlation between speakers' use of *be like* in their L2 and *être comme* in their L1, we isolated thirteen younger (19- to 25-year-old) speakers who made use of both variants.[10] Figure 2 shows the proportion of quotative use comprising *être comme* and *be like* in individual speakers' L1 and L2 respectively. Although there are some discrepancies in individual rates of *être comme* and *be like* (see, e.g., speakers 002, 008, 025), for other speakers, L1 and L2 rates do not diverge markedly. Indeed, a Pearson's Correlation reveals that in the aggregate, there is a significant moderate positive

---

[10]Of the fourteen speakers in this age cohort, one produced a limited number of quotatives and no occurrences of *être comme*. This speaker was excluded from the comparison of *être comme* and *be like*.

Table 9. Distribution of quotative variants by speaker sex and age in the L1 Corpus of French

| | 19–25 | | | | 50–77 | |
| | Females | | Males | | Females | |
Variant	N	%	N	%	N	%
Être comme	178	60	77	39	6	3
Dire	35	12	43	22	124	62
Discourse *comme*	21	7	14	7	2	1
C'est comme	26	9	16	8	5	2
Zero	8	3	15	8	39	20
Se dire	4	1	3	2	7	4
Faire	0	0	4	2	2	1
Other	27	9	25	13	15	8
Total	299		197		200	
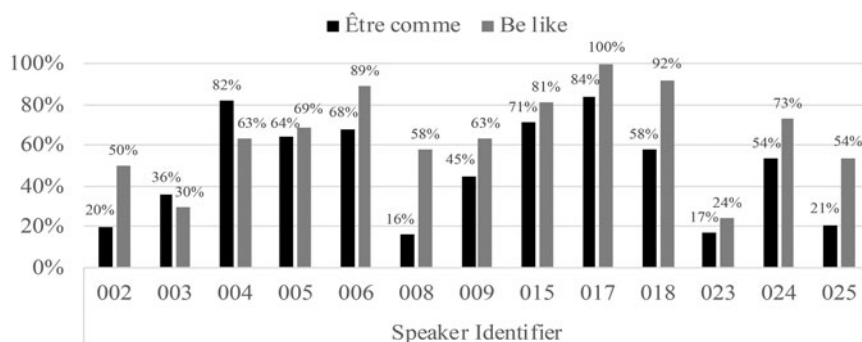
Figure 2. Proportion of individual quotative use comprising *être comme* and *be like* in speakers' L1 and L2 respectively

association between *être comme* and *be like*, $r(11) = .75$, $p = .003$. One tentative conclusion that we draw from this correlation is that speakers who tend to be progressive in the use of *être comme* in their L1 are likewise liberal in their use of *be like* in their L2. By the same token, speakers who are conservative in their use of *être comme* tend to be more restrained in their use of *be like*. Prior to making more definitive claims about these associations, we caution that our results, while suggestive, must be tempered by the small number of speakers ($N = 13$) on which they are based.

A more exacting test of possible transfer effects is one which transcends reliance on superficial form-based correspondences and their attendant usage rates in favour of detailed comparisons of the probabilistic constraints on variant selection in the TL, the L2 as well as L2 speakers' L1. We now turn to this three-way comparison, with a view to ascertaining the extent to which form-based similarities and commensurate rates of variant occurrence can additionally be equated with the functional isomorphy of variant forms.

7.4.2 Conditioning of variant forms

Table 10 shows the output of a mixed-effects regression analysis of the contribution of social and linguistic factors to the choice of *be like* in the TL benchmark variety using Rbrul (Johnson 2009). In this and ensuing tables of results, individual speaker is run as a random effect to control for any individual biases that may arise as a result of idiosyncratic patterns of variation (Johnson 2009).

The numerical formalisms associated with regression analysis are to be interpreted as follows. The input probability is a measure of the overall probability that *be like* (or *être comme*) will occur in the dataset in question. The log likelihood indicates the goodness of fit of the regression model to the dataset under consideration and the R^2 value represents the proportion of variation explained by the model. Key to the interpretation of

Table 10. Rbrul mixed-effects logistic regression analysis of the contribution of social and linguistic factors to the selection of *be like* in the speech of 19- to 28-year-olds in the Target Language Corpus of English (OEC)[11]

Input probability	0.552			
Log likelihood	−330.618			
R2 total	0.657			
N	586/911			
Grammatical person (p < 1.92e-31)	Log Odds	Centred FW	%	N
1st Person	1.529	0.822	82	294
3rd Person	0.988	0.729	70	446
Other	−2.517	0.075	17.5	171
Tense (p > 2.86e-36)				
Conversational Historical Present	1.500	0.818	87	244
Past tense	0.325	0.580	72	355
Present tense	0.079	0.520	54	121
Other	−1.904	0.130	28	191
Mimesis (p < 0.0007)				
Mimetic	0.414	0.602	72	448
Non-Mimetic	−0.414	0.398	57	463
Content of quote (n.s.)				
Reported thought	0.261	0.565	72	202
Reported speech	−0.216	0.446	62	690
Speaker sex (n.s.)				
Female	0.645	0.656	72	539
Male	−0.645	0.344	53	372

[11]For content of the quote, ambiguous tokens (i.e., ones where it was not possible to distinguish reported thought from speech) are not reported here.

contextual effects on variant choice is the hierarchy of constraints, represented by the log-odds (LO) and the centred factors weights (FW). Log-odds with a positive value indicate that a factor listed on the left-hand side of the table has a favouring effect on variant choice, whereas those with a negative value have a disfavouring effect. The centred factor weights have a similar interpretation: factor weights above 0.5 favour the application value (i.e., *be like* in Table 10), whereas those below 0.5 disfavour the application value.

Here and in subsequent tables of results, we rely on two lines of evidence in interpreting the output of the regression analysis: statistical significance and the constraint hierarchy, or the ordering of individual constraints in each independent factor group or predictor. Together, this information offers a detailed window on the quantitative conditioning of variant choice.

What key structural information can be abstracted from the results shown in Table 10? Grammatical person, tense and mimesis all make statistically significant contributions to the selection of *be like* and the direction of effect for each predictor is aligned with the results of previous investigations targeting quotative variation and change: *be like* is preferentially selected with first person contexts, the conversational historical present as well as with quoted extracts containing mimetic effects (see Ferrara and Bell 1995, Buchstaller and D'Arcy 2009, Buchstaller 2014).

Although the direction of effects for content of the quote and speaker sex is also aligned with results documented in earlier studies, neither predictor is selected as significant. Buchstaller (2014) reports that the effect of speaker sex is labile across communities and across time. Although previous investigations targeting Canadian English (Tagliamonte and D'Arcy 2004) have noted a strengthening of this effect, along with an increase in the overall use of *be like,* it is not improbable that this effect should weaken as *be like* saturates the quotative system. Attenuation of the effect of content of the quote is likewise not unprecedented: Hunt Gardner et al. (2021) observe that, as *be like* swamps the quotative system of Toronto English, rates of *be like* used with reported dialogue and thought begin to converge.

How do the constraint hierarchies in the L2 data compare with their counterparts in the TL baseline variety? Table 11 shows the results of the regression analysis for younger (19- to 33-year-olds) in the L2 sample. The ranking of constraints governing variant choice shares a number of parallels with those in the TL benchmark, although two predictors, content of the quote and speaker sex, return significant effects in the L2 dataset but not in the TL baseline. Tense is the only category in which the constraint rankings differ in the comparison varieties. In the L2 data, *be like* is only favoured with the conversational historical present, a correlation that is also replicated, as observed earlier, in the TL benchmark variety. But, in contrast with the L2 dataset, *be like* is moderately favoured in the TL baseline with the past tense, and additionally exhibits a very weak association with the present tense.

Of particular importance in the L2 results is the significant contribution of English language proficiency, as measured by the CEPI, to the selection of *be like*, with mid-high/high proficiency speakers exhibiting a greater propensity to select this variant than less proficient speakers, as we initially hypothesized.

The results in Table 11 raise the possibility that proficiency not only correlates with differential usage rates of *be like* (cf. Table 8) but is also associated with underlying structural differentiation in the conditioning of variant choice. To address this

Table 11. Rbrul mixed-effects logistic regression analysis of the contribution of social and linguistic factors to the selection of *be like* in the speech of 19-33-year-olds in the L2 Corpus of English[12]

Input probability	0.714			
Log likelihood	−201.879			
R2 total	0.628			
N	347/535			
Grammatical person (p < 1.36e-22)	Log Odds	Centred FW	%	N
1st Person	1.412	0.804	89	198
3rd Person	0.678	0.633	76	181
Other	−2.090	0.110	21	156
Tense (p < 0.0008)				
Conversational Historical Present	1.394	0.801	93	71
Past tense	−0.110	0.472	78	156
Present tense	−0.461	0.387	67	175
Other	−0.823	0.305	32	133
Mimesis (p < 0.0142)				
Mimetic	0.349	0.586	72	276
Non-Mimetic	−0.349	0.414	57.5	259
Content of quote (p < 0.0003)				
Reported thought	0.602	0.646	67	178
Reported speech	−0.602	0.354	64	357
Speaker sex (p < 0.0331)				
Female	0.511	0.625	75	327
Male	−0.511	0.375	49.5	208
CEPI (p < 0.004)				
Mid-high/high	0.738	0.677	73	378
Low/mid-low	−0.738	0.323	47	157

possibility, we redid the analysis of the L2 data, focusing exclusively on the mid-high/ high proficiency speakers.[13]

---

[12]One young L2 speaker, who produced no tokens of *be like*, is omitted from the regression analysis, hence the Total N=535 and not 541, as inferred from Table 7 (see 19-33 females and males). Although the percentage values for reported thought (67%) and speech (64%) are similar, this predictor is nonetheless selected as significant in the mixed-effects regression analysis. It is also significant when run alone with the dependent variable, along with speaker as a random effect.

[13]The low/mid-low proficiency group yielded 157 quotatives only. We can expect reduced statistical power in datasets containing a limited number of tokens. In fact, most predictors in the low/mid-low proficiency cohort failed to return significant effects. Suffice it to mention here that, with the exception of tense and speaker sex, the constraint hierarchies for the low/mid-low proficiency group largely resemble what we see in the logistic regression analysis of the data associated with the mid-high/higher proficiency speakers.

Apart from the fact that mimesis is selected as significant in the TL baseline depicted in Table 10, but not in the L2 cohort shown in Table 12, the few remaining discrepancies between the TL and L2 varieties in the conditioning of variant choice (e.g., in the respective contribution of the present tense to variant selection) are hardly of the magnitude that would warrant any claim of substantial underlying structural divergence between the comparison varieties. The very minor differences in conditioning that distinguish the advanced L2 and TL speaker cohorts are quantitative rather than qualitative in nature. As such, we venture that they would be all but invisible to native speakers of the TL variety. These findings, taken together with similar variant inventories and usage rates docu-

Table 12. Rbrul mixed-effects logistic regression analysis of the contribution of social and linguistic factors to the selection of *be like* in the speech of 19–33-year-olds in the L2 Corpus of English (mid-high/high proficiency speakers only)[14]

Input probability	0.834			
Log likelihood	−132.236			
R2 total	0.533			
N	274/378			
Grammatical person (P < 1.63e-14)	Log Odds	Centred FW	%	N
1st Person	1.345	0.793	94	139
3rd Person	0.533	0.630	87	135
Other	−1.877	0.133	25	104
Tense (p < 0.0028)				
Conversational Historical Present	1.202	0.769	94	54
Past tense	0.379	0.594	93	102
Present tense	−0.454	0.388	72	133
Other	−1.126	0.245	36	89
Mimesis (n.s)				
Mimetic	0.283	0.570	77	208
Non-Mimetic	−0.283	0.430	67	170
Content of quote (n.s.)				
Reported thought	0.427	0.605	68	128
Reported speech	−0.427	0.395	75	250
Speaker sex (n.s.)				
Female	0.464	0.614	79	277
Male	−0.464	0.386	55	101

[14]Inspection of the results for quote content reveals that the hierarchy of percentage values associated with reported thought and speech respectively is the inverse of the corresponding hierarchy of centred factor weights and log-odds. This is not the result of any interaction with other predictors included in the same model, as exactly the same result is obtained when content is run alone with the dependent variable, together with speaker as a random effect.

mented in both the TL and higher proficiency L2 cohorts, lead us to infer that the latter are *bona fide* participants in ongoing change affecting the TL baseline variety.

To what extent might the ability of L2 speakers to participate in ongoing change in the TL be abetted by the environmental constraints operating on *être comme* in their native French? Table 13 shows the results of a mixed-effects regression analysis depicting the contribution of social and linguistic predictors to the selection of *être comme* in the speech of 19- to 25-year-old francophones. A first point of difference between Tables 12 and 13 is that all the independent linguistic variables in the speakers' L1 make a significant contribution to *être comme*, whereas only grammatical person and tense exhibit significant effects on the selection of *be like* in the speech of

Table 13. Rbrul mixed-effects logistic regression analysis of the contribution of social and linguistic factors to the selection of *être comme* in the speech of 19- to 25-year-olds in the L1 Corpus of French[15]

Input probability	0.470			
Log likelihood	−155.115			
R2 total	0.774			
N	255/488			
Grammatical person (p < 1.91e-21)	Log Odds	Centred FW	%	N
1st Person	1.789	0.857	81	190
3rd Person	0.476	0.617	55	153
Other	−2.265	0.094	12	145
Tense (p < 2.73e-19)				
Conversational Historical Present	1.521	0.821	77	96
Imperfect tense	1.234	0.775	80	114
Present tense	0.212	0.553	56	150
Other	−2.967	0.049	5	128
Mimesis (p < 0.0002)				
Mimetic	0.596	0.645	61	259
Non-Mimetic	−0.596	0.355	43	229
Content of quote (p < 0.0193)				
Reported thought	0.457	0.612	61	192
Reported speech	−0.457	0.388	47	296
Speaker sex (n.s.)				
Female	0.608	0.648	61	291
Male	−0.608	0.352	39	197

---

[15]The Total N=488, and not 496, because one speaker was excluded on the grounds that she produced no tokens of *être comme*. Although the percentage values for females and males are noticeably distinct, this factor group is selected as non-significant. According to Johnson (2009: 365), the inclusion of a speaker random effect ensures that external predictors will achieve significance only "when they are strong enough to rise above […] inter-speaker variation."

higher-proficiency L2 speakers. In terms of the constraint ranking for tense, the present tense favours *être comme* in Table 13 but has a corresponding disfavour effect on *be like* in Table 12. Notwithstanding what again amounts to minor differences, the detailed contextual conditioning of *être comme* in the L1 data evinces compelling affinities with the constraints governing *be like* in the L2 data. The existence of such parallelisms bolsters the possibility of L1 transfer effects on speakers' L2.

8. Discussion

In this article, we have described the construction of three new corpora designed to leverage novel insights into L2 acquisition by comparing L2 speech, a community-based TL baseline variety, and L2 speakers' first language, French. Exemplifying with a case study addressing the L2 acquisition of quotative variation and change in the local TL benchmark variety, we targeted the precipitous rise of the *be like* quotative variant, reported to be a flagship case of change in the English-speaking world. We additionally factored into our analysis a parallel development in L2 speakers' native French involving the ascendancy of *être comme*, spearheaded by francophone youth. Our diagnostic measures of the L2 acquisition of change focused, in ascending order of importance, on the replication of the TL inventory of quotative variants, approximation of TL variant usage rates, and, crucially, the reproduction of the TL set of implicit constraints governing the selection of *be like*.

Perhaps the most surprising result of this enterprise is the key revelation that the candidate for change in the TL, *be like*, is not only similarly distributed in the TL baseline variety and the discourse of mid-high/high proficiency L2 speakers, it is also largely sensitive in L2 discourse to the same set of probabilistic constraints that govern its selection in the TL. This unexpected finding not only runs counter to claims that variables at the syntax-discourse interface should be challenging to acquire (Sorace 2004), it is also at variance with the thesis that sociolinguistic constraints on variant use should transfer from TL to L2 speakers in part rather than in toto (Schleef et al. 2011). Our findings are all the more remarkable as they resonate with Davydova's (2021) conclusions based on independent datasets and distinct L2 populations. In contrast with that study, however, our results additionally indicated that rates of *be like* are broadly comparable in the higher-proficiency L2 and TL speaker groups.

We cannot rule out the possibility that our approach has overlooked, or failed to detect, major differences between the TL and L2 comparison varieties pertaining to the acquisition of the targeted change in progress, perhaps by virtue of the relatively small number of independent social and linguistic variables incorporated into our analysis.[16] Nevertheless, our tools were specifically configured to detect gradient differences between the comparison varieties, affording a level of analytical granularity capable of revealing even very subtle inter-varietal distinctions. This leads us to stress

---

[16]A reviewer raised the possibility that speakers (N = 9) residing on the Quebec side of the provincial border might exhibit different L2 patterns vis-à-vis their counterparts in Ontario. Of the six younger (19-25) L2 speakers residing in Quebec, two have low CEPI scores and produced limited occurrences of *be like*. The other four, however, have mid-high CEPI scores and exhibited individual rates of *be like* ranging from a low of 53% to a high of 73% of all quotatives produced. The same four speakers report many anglophones in their personal social networks.

once again that, to the extent we detected differences between higher-proficiency L2 speakers and their TL counterparts, such differences were relatively trivial in nature and would not constitute strong evidence of partial acquisition.

How, then, can we account for the detailed approximations of higher-proficiency L2 speakers of English to TL vernacular norms when current theorizing and the weight of previous sociolinguistic research on L2 acquisition are at odds with such an outcome? At least part of the answer resides in the constellation of favourable social circumstances that are hospitable to high levels of L2 attainment in the community we investigated. Specifically, in the higher-proficiency L2 population that we examined, more than half the speakers stated that native anglophones made up 50% or more of their personal social networks, bolstering our working hypothesis that those speakers with the strongest personal ties to the local TL community should be expected to make the most substantial gains in acquiring native-like sociolinguistic competence in the TL. This expectation is congruent with recurrent claims in the literature that high levels of exposure to, and engagement with, TL norms is a crucial predictor of the L2 mastery of TL vernacular features (Sankoff et al. 1997, Blondeau et al. 2002, Howard et al. 2013).

Yet social factors, no matter how conducive they may be to high levels of L2 attainment, cannot be wholly determinative of the findings we have uncovered here. If that were the case, then the very same set of favourable social circumstances (protracted contact between English and French, extensive bilingualism, etc.), widely believed to be major drivers of convergent change (Thomason 2001), together with the extensive structural similarities between English and French, would be expected to result in promiscuous contact-induced effects in the minority French examined here, given its intense contact with majority English. Yet decades of empirical research in the Canadian National Capital Region, examining a wide array of grammatical variables, have failed to turn up robust evidence in support of such a scenario (Poplack and Levey 2010).

This forces us to seek additional explanations for our results in the specific details of the variable we targeted for analysis. As previously mentioned, Davydova (2021) builds a solid case for the socio-cognitive salience of *be like* in discourse, with a conspiracy of factors facilitating L2 speakers' acquisition of probabilistic patterns of structured variation. We note in the TL variety examined here that the sheer quantitative preponderance of *be like* in the quotative system of contemporary Canadian anglophone youth effectively renders this construction the default variant of choice. In that variety, the fact that *be like* has achieved hegemonic status and is not (or no longer) embedded in a system where several different options *vigorously* compete with one other to perform similar grammatical functions is, as Davydova (2021) mentions, an important factor that renders its acquisition more manageable by L2 speakers. And we know from cross-linguistic research that unmarked and less tightly integrated elements are reported to be particularly susceptible to inter-linguistic influence (Thomason 2001).

Perhaps the key to explaining our results resides in the fact that L2 speakers' native French hosts a counterpart feature to English *be like*, which, although comparatively less frequent than its English congener, is nonetheless predominant in the speech of contemporary francophone youth, as our distributional analysis confirmed. This fact,

together with our demonstration that *être comme* essentially occupies the same functional niches as its English counterpart, confers a distinct advantage on proficient L2 speakers' capacity to replicate the structural details of an ongoing, and, to all intents and purposes, parallel change in the TL. While we reserve judgment on whether shared structural and functional parallelisms between *être comme* and *be like* uniquely enhance the possibility of forward transfer effects (i.e., L1 influence on L2), as opposed to also operating bidirectionally (i.e., L2 influence on L1; see Kartushina et al. 2016), we submit that such "interlingual identifications" (Weinreich 1968: 7), and, in particular, the underlying environmental constraints that condition them, can subtly refine and enhance our understanding of levels of L2 attainment.

Our results add to the small body of research indicating that native-likeness in L2 acquisition, at least with respect to the linguistic change we investigated, may not be so rare as to qualify it as tangential to L2 acquisition theory (see e.g. Birdsong 2004). While we by no means interpret the results of our case study as a blanket refutation of the widely espoused hypothesis that the L2 acquisition of TL features will entail some weakening, re-organization or re-interpretation of original TL constraints and patterns (Labov 2007, Schleef et al. 2011), we would nonetheless argue that, on occasion, a felicitous combination of social and structural factors, coupled with advanced L2 proficiency, may exceptionally mitigate the likelihood of those outcomes.

We conclude by emphasizing that such special cases elude casual observation and can only be brought to light via systematic comparative analysis, capitalizing on the structured heterogeneity found in community-based speech varieties. The corpora we have built enabling such an analysis are uniquely positioned for extending this line of inquiry to other variables situated in other areas of L2 grammar.

References

Adamson, Hugh Douglas, and Vera Regan. 1991. The acquisition of community speech norms by Asian immigrants learning English as a second language: A preliminary study. *Studies in Second Language Acquisition* **13**: 1–22.

Backus, Ad. 2005. Code-switching and language change: One thing leads to another? *International Journal of Bilingualism* **9**: 307–340.

Baker, Colin. 1992. *Attitudes and language*. Clevedon: Multilingual Matters.

Bayley, Robert, and Vera Regan. 2004. Introduction: The acquisition of sociolinguistic competence. In *Journal of Sociolinguistics* **8**(3): *Acquisition of sociolinguistic competence*, ed. Robert Bayley and Vera Regan, 323–338.

Bayley, Robert, and Dennis Preston. 2008. Variation and second language grammars. *Studies in Hispanic and Lusophone Linguistics* **1**: 385–397.

Bayley, Robert, and Chelsea Escalante. 2022. Variationist approaches to second language acquisition. In *The Routledge handbook of second language acquisition and sociolinguistics,* ed. Kimberly L. Geeslin, 3–16. New York: Routledge.

Birdsong, David. 2004. Second language acquisition and ultimate attainment. In *The handbook of applied linguistics,* ed. Alan Davies and Catherine Elder, 82–105. Oxford: Blackwell.

Birdsong, David. 2006. Age and second language acquisition and processing: A selective overview. *Language Learning* **56**: 9–49.

Blondeau, Hélène, Naomi Nagy, Gillian Sankoff, and Pierrette Thibault. 2002. La couleur locale du français L2 des anglo-montréalais. *Acquisition et interaction en langue étrangère* **17**: 1–21.

Blondeau, Hélène, Nathalie Dion, and Zoe Ziliak Michel. 2014. Future temporal reference in the bilingual repertoire of Anglo-Montrealers: A twin variable. *International Journal of Bilingualism* **18**: 674–692.

Bourhis, Richard. 1984. Ethnic and language attitudes in Quebec. In *Ethnicity and culture in Canada: The research landscape*, ed. John W. Berry and Jean A. Laponce, 322–360. Toronto: University of Toronto Press.

Buchstaller, Isabelle. 2006. Diagnostics of age-graded linguistic behaviour: The case of the quotative system. *Journal of Sociolinguistics* **10**(1): 3–30.

Buchstaller, Isabelle. 2014. *Quotatives: New trends and sociolinguistic implications*. Oxford: Wiley Blackwell.

Buchstaller, Isabelle, and Alexandra D'Arcy. 2009. Localised globalisation: A multilocal, multi-variate investigation of *be like*. *Journal of Sociolinguistics* **15**: 291–331.

Buchstaller, Isabelle, and Ingrid van Alphen. 2012. Introductory remarks on new and old quotatives. In *Quotatives: Cross-linguistic and cross-disciplinary perspectives,* ed. Isabelle Buchstaller and Ingrid van Alphen, xi-xxx. Amsterdam: John Benjamins.

Butters, Ronald. 1982. Editor's note [on 'be+like']. *American Speech* **57**: 149.

Choi, Jihyun Karen, and Chloé Diskin-Holdaway. 2022. The acquisition of quotatives and quotative *be like* among Chinese L2 speakers of English in Australia. *Languages* **7**(2): 123. https://doi.org/10.3390/languages7020123

Corrigan, Karen. 2020. *Linguistic communities and migratory processes: Newcomers acquiring sociolinguistic variation in Northern Ireland*. Berlin: De Gruyter Mouton.

Davydova, Julia. 2021. The role of sociocognitive salience in the acquisition of structured variation and linguistic diffusion: Evidence from quotative *be like*. *Language in Society* **50**: 171–196.

Deng, Delin. 2023. *"She's like why you speak English while dreaming?"*: A corpus-based study of quotative markers used by Chinese speakers of L2 English. *Languages* **8**(1): 51. https://doi.org/10.3390/languages8010051

Dion, Nathalie, and Shana Poplack. 2007. Linguistic mythbusting: The role of the media in diffusing change. Paper presented at NWAV 36, Philadelphia, University of Pennsylvania, October 11–14.

Diskin, Chloé, and Stephen Levey. 2019. Going global and sounding local: Quotative variation and change in L1 and L2 speakers of Irish (Dublin) English. *English World-Wide* **40**(1): 53–78.

Dörnyei, Zoltán, Kata Csizér, and Nóra Németh, 2006. *Motivation, language attitudes and globalisation: A Hungarian perspective*. Clevedon, England: Multilingual Matters.

Durham, Mercedes, Bill Haddican, Eytan Zweig, Daniel Ezra Johnson, Zipporah Baker, David Cokeram, Esther Danks, and Louise Tyler. 2012. Constant linguistic effects in the diffusion of *be like*. *Journal of English Linguistics* **40**(4): 316–337.

Ferrara, Kathleen, and Barbara Bell. 1995. Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of be + like. *American Speech* **70**: 265–289.

Gardner, Robert, C. 1985. *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, Robert, C., and Wallace E. Lambert. 1959. Motivational variables in second-language acquisition. *Canadian Journal of Psychology* **13**: 266–272.

Gardner, Robert C., and Wallace E. Lambert. 1972. *Attitudes and motivation in language learning*. Rowley, MA: Newbury House.

Geeslin, Kimberley, and Avizia Yim Long. 2014. *Sociolinguistics and second language acquisition: Learning to use language in context*. London: Routledge

Giles, Howard, Richard Bourhis, and Donald M. Taylor. 1977. Towards a theory of language in ethnic group relations. In *Language, ethnicity and intergroup relations,* ed. Howard Giles, 307–348. London: Academic Press.

Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.

Heine, Bernd, and Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press

Howard, Martin, Raymond Mougeon, and Jean-Marc Dewaele. 2013. Sociolinguistics and second language acquisition. In *The Oxford handbook of sociolinguistics,* ed. Robert Bayley, Richard Cameron, and Ceil Lucas, 340–359. Oxford: Oxford University Press.

Hunt Gardner, Matt, Derek Denis, Marisa Brook, and Sali A. Tagliamonte. 2021. *Be like* and the constant rate effect: From the bottom to the top of the S-curve. *English Language and Linguistics* **25**: 281–324.

Johnson, Daniel Ezra. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3: 359–383.

Kartushina, Natalia, Ulrich H. Frauenfelder, and Narly Golestani. 2016. How and when does the second language influence the production of native speech sounds: A literature review. *Language Learning* 66: 155–186

Kircher, Ruth. 2009. Language attitudes in Quebec: A contemporary perspective. Doctoral dissertation, Queen Mary, University of London.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1984. Field methods of the project on linguistic change and variation. In *Language in use: Readings in sociolinguistics,* ed. John Baugh and Joel Sherzer, 28–54. Englewood Cliffs, NJ: Prentice-Hall.

Labov, William. 2007. Transmission and diffusion. *Language* 83: 344–387.

Levey, Stephen, Karine Groulx, and Joseph Roy. 2013. A variationist perspective on discourse-pragmatic change in a contact setting. *Language Variation and Change* 25: 225–251.

Magliacane, Annarita, and Martin Howard. 2019. The role of learner status in the acquisition of pragmatic markers during study abroad: The use of 'like' in L2 English. *Journal of Pragmatics* 146: 72–86.

Matras, Yaron, and Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31: 829–865.

Meyerhoff, Miriam, and Erik Schleef. 2014. Hitting an Edinburgh target: Immigrant adolescents' acquisition of variation in Edinburgh English. In *Sociolinguistics in Scotland,* ed. Robert Lawson, 103–128. Basingstoke: Palgrave Macmillan.

Mougeon, Raymond, Terry Nadasdi, and Katherine Rehner. 2004. The learning of spoken French variation by immersion students from Toronto. *Journal of Sociolinguistics* 8: 408–432.

Mougeon, Raymond, Terry Nadasdi, and Katherine Rehner. 2010. *The sociolinguistic competence of immersion students*. Bristol: Multilingual Matters.

Nagy, Naomi, Hélène Blondeau, and Julie Auger. 2003. Second language acquisition and 'real' French: An investigation of subject doubling in the French of Montreal Anglophones. *Language Variation and Change* 15: 73–103.

Nagy, Naomi, Christine Moisset, and Gillian Sankoff. 1996. On the acquisition of variable phonology in L2. *University of Pennsylvania Working Papers in Linguistics* 3: 111–126.

Odlin, Terence. 2003. Cross-linguistic influence. In *The Handbook of second language acquisition,* ed. Catherine J. Doughty and Michael H. Long, 436–486. Oxford: Blackwell.

Papen, Robert. 2022. French and English in contact in North America. In *The Cambridge handbook of language contact, Volume 1: Population movement and language change*, ed. Salikoko Mufwene, and Anna María Escobar, 505–539. Cambridge: Cambridge University Press.

Poplack, Shana. 1989. The care and handling of a mega-corpus. In *Language variation and change,* ed. Ralph Fasold and Deborah Schiffrin, 411–451. Amsterdam: Benjamins.

Poplack, Shana. 2018. *Borrowing: Loanwords in the speech community and in the grammar*. Oxford: Oxford University Press.

Poplack, Shana, Nathalie Dion, Suzanne Robillard, and Basile Roussel. 2019. Language mixing causes a stir: Confronting science and public opinion. In *Linguistic duality, de jure and de facto: Marking the 50th anniversary of the Official Languages Act,* ed. Jack Jedwab, 53–57. Montreal: Association for Canadian Studies.

Poplack, Shana, Allison Lealess, and Nathalie Dion. 2013. The evolving grammar of the French subjunctive. *Probus* 25(1): 139–195.

Poplack, Shana, and Stephen Levey. 2010. Contact-induced grammatical change: A cautionary tale. In *Language and space—An international handbook of linguistic variation. Vol. 1: Theories and methods,* ed. Peter Auer and Jürgen Erich Schmidt, 391–419. Berlin: Mouton de Gruyter.

Poplack, Shana, and Marjorie Meechan. 1998. How languages fit together in code-mixing. *International Journal of Bilingualism* 2: 127–138.

Poplack, Shana, and Sali Tagliamonte. 2001. *African American English in the diaspora*. Oxford: Blacwkell.

Poplack, Shana, James A. Walker, and Rebecca Malcolmson. 2006. An English "like no other"?: Language contact and change in Quebec. *Canadian Journal of Linguistics* 51(2/3): 185–213.

Regan, Vera, Martin Howard, and Isabelle Lemée. 2009. *The acquisition of sociolinguistic competence in a study abroad context*. Bristol: Multilingual Matters.

Rehner, Katherine, and Raymond Mougeon. 2022. Variationist methods of analysis. In *The Routledge handbook of second language acquisition and sociolinguistics,* ed. Kimberly Geeslin, 200–211. London: Routledge.

Rehner, Katherine, Raymond Mougeon, and Terry Nadasdi. 2003. The learning of sociolinguistic variation by advanced FSL learners: The case of *nous* versus *on* in immersion French. *Studies in Second Language Acquisition* **25**: 127–156.

Romaine, Suzanne. 1995. *Bilingualism*. 2nd ed. Oxford: Blackwell.

Sankoff, Gillian, Pierrette Thibault, Naomi Nagy, Hélène Blondeau, Marie-Odile Fonollosa, and Lucie Gagnon. 1997. Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* **9**: 191–218.

Schleef, Erik. 2017. Developmental sociolinguistics and the acquisition of T-glottalling by immigrant teenagers in London. In *Acquiring sociolinguistic variation,* ed. Gunther de Vogelaer and Matthias Katerbow, 311–347. Amsterdam: John Benjamins.

Schleef, Erik, Miriam Meyerhoff, and Lynn Clark. 2011. Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London. *English World-Wide* **32**(2): 206–236.

Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics* **10**: 209–231.

Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition* **7**: 143–5.

Statistics Canada. 2021. *Census of population: Ottawa-Gatineau census metropolitan area.* https://www.2.statcan.gc.ca/census-recensement/2021/

Tagliamonte, Sali, and Alexandra D'Arcy. 2004. 'He's like, she's like:' The quotative system in Canadian youth. *Journal of Sociolinguistics* **8**: 493–514.

Tagliamonte, Sali, and Alexandra D'Arcy. 2007. Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change* **19**: 199–217.

Tagliamonte, Sali, and Rachel Hudson. 1999. *Be like* et al. beyond America: the quotative system in British and Canadian youth. *Journal of Sociolinguistics* **3**(2): 147–172.

Tagliamonte, Sali, Alexandra D'Arcy, and Celeste Rodríguez Louro. 2016. Outliers, impact, and rationalization in linguistic change. *Language* **92**: 824–849.

Tasseva-Kurktchieva, Mila. 2015. Can production precede comprehension in L2 acquisition? *Second Language Research* **31**: 493–522.

Taylor, Donald M., Roch Maynard, and Elizabeth Rhéault. 1977. Threat to ethnic identity and second-language learning. In *Language, ethnicity, and intergroup relations,* ed. Howard Giles, 98–118. London: Academic Press.

Thomason, Sarah. G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah. G. 2013. Innovation and contact: the role of adults (and children). In *English as a contact language,* ed. Daniel Schreier and Marianne Hundt, 283–297. Cambridge: Cambridge University Press.

Torres Cacoullos, Rena, and Catherine E. Travis. 2018. *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge: Cambridge University Press.

Trudgill, Peter. 2014. Diffusion, drift, and the irrelevance of media influence. *Journal of Sociolinguistics* **18**(2): 213–222.

Uritescu, Dorin, Raymond Mougeon, Katherine Rehner, and Terry Nadasdi. 2004. Acquisition of internal and external constraints of variable schwa deletion by French immersion students. *International Review of Applied Linguistics* **42**: 349–364.

Vincent, Diane. 2005. The journey of non-standard discourse markers in Quebec French. *Journal of Historical Pragmatics* **6**: 188–210.

Weinreich, Uriel. 1968. *Languages in contact: Findings and problems*. The Hague: Mouton [1953].

Weinreich, Uriel, William Labov and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for historical linguistics,* ed. Winfred Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.