REVISITING PHONETIC INTEGRATION IN BILINGUAL BORROWING

Shana Poplack

University of Ottawa

Nathalie Dion

University of Ottawa

Suzanne Robillard

University of Ottawa

John C. Paolillo

Indiana University Bloomington

This article investigates whether speakers marshal phonetic integration as a strategy to distinguish language-contact phenomena. Systematic comparison of the behavior of individuals, diagnostics, and language-mixing types (code-switches, established loanwords, and nonce borrowings) reveals variability at every level of the adaptation process, providing strong evidence that bilinguals do not phonetically distinguish other-language words, nonce or dictionary-attested, in a uniform way. This is in striking contrast to the community-wide morphosyntactic treatment they afford this same material when borrowing it: immediate, quasi-categorical, and consistent. This confirms that phonetic and morphosyntactic integration are independent. Only the latter is a reliable metric for distinguishing language-mixing types.*

Keywords: metrics for distinguishing language-contact phenomena, phonetic integration, lexical borrowing, loanword adaptation, nonce borrowings, code-switches, loanword phonology

1. Introduction. Research on spontaneous language mixing has shown that bilinguals tend overwhelmingly to imbue the words they borrow with the morphological and syntactic features of the recipient language, and that such INTEGRATION occurs well before the borrowings have achieved the status of bona fide loanwords (e.g. Adalar & Tagliamonte 1998, Budzhak-Jones 1998, Eze 1997, Ghafar Samar & Meechan 1998, Poplack 2018, Poplack & Dion 2012, Poplack & Meechan 1998a, Turpin 1995). Indeed, where the extralinguistic characteristics of loanword status (frequency of use and diffusion across the community) are lacking, this quasi-immediate adoption of the grammatical apparatus of the recipient language is the principal basis on which we can identify loanwords as borrowed. At the same time, these grammatically integrated items often seem to resist integration at the phonetic/phonological level. This is at odds with the widespread assumption that loanwords adopt all of the features of the language into which they are incorporated. Where variability has been noted, it has been viewed as either transitional or exceptional in some other sense (e.g. Chang 2008, Crawford 2007, Heffernan 2007, Kang 2010, Kenstowicz & Suchato 2006, LaCharité & Paradis 2005, Paradis & LaCharité 1997, 2008). As a result, other-language items with other-language phonetic realizations are often considered not to be loanwords.

As currently conceived by many phonologists, loanword ADAPTATION—the means by which borrowed words are transformed to fit the phonetic and phonological system of a borrowing, or recipient, language (L_R)—proceeds roughly as follows: L_R receives input from a donor language (L_D), replete with the characteristic features of that language. If those features happen to be 'illicit' in L_R, they must be adapted, by means of 'repair strategies', operations that ensure conformity with the constraints of L_R (e.g. Paradis & LaCharité 1997). Such repairs are thought to apply predictably and categorically, although consensus has yet to be achieved over precisely how they arise or how

they are implemented (e.g. through perceptual means or via category-wise phonological representations). To complicate matters, while accommodating to the phonological restrictions of L_R, the borrower's purported goal is to remain as faithful as possible to the source word (e.g. Kenstowicz & Suchato 2006), so as to achieve the best match between them. Perhaps not surprisingly, a number of 'puzzling patterns' (Kang 2011) of loanword adaptation have been identified. Among those cited by Kang are 'divergent repair' (Kenstowicz 2005), wherein the type of adaptation applied to the loanword contradicts the native L_R repair strategy, 'unnecessary repair' (Peperkamp 2005), wherein a structure that is licit in L_R is repaired nonetheless, and 'failure to repair' (Kang's 'differential importation'), as when an illicit segment in L_R is tolerated in some, but not all, loanwords.

Much research has been devoted to explaining these unruly facts, and has even led to the creation of an entire field, LOANWORD PHONOLOGY. Among the major issues that have preoccupied its practitioners are: (i) whether the type of information referred to by the borrower in searching for a licit L_R counterpart to an illicit L_D form is phonetic (Lev-Ari et al. 2014, Peperkamp 2005, Peperkamp & Dupoux 2003, Peperkamp et al. 2008, Silverman 1992), phonological (Jacobs & Gussenhoven 2000, Kenstowicz & Suchato 2006, LaCharité & Paradis 2005, Paradis & LaCharité 1997, 2008, 2012), or both (Chang 2008, Kang 2011, Yip 2006), and by extension, whether the adaptation process takes place in perception or production, and (ii) whether certain kinds of match are preferred over others. Opinions remain deeply divided, with little prospect of imminent resolution. Yet despite the lack of consensus, there seems to be widespread agreement on at least one issue: speakers converge on a shared adaptation strategy for their L_D-origin words (e.g. Boersma & Hamann 2009, Broselow 2009, Calabrese 2009, Crawford 2007, Kang 2010, Kenstowicz & Suchato 2006, Paradis & LaCharité 1997, 2008), and this strategy, in turn, is viewed as quasi-invariant. According to Paradis and LaCharité (2008:121), for example, 'the adaptations found in loanwords are overwhelmingly categorical, i.e. there is generally an exclusive or strongly dominant strategy for adapting an unacceptable L_2 phoneme'.[1] In this article, we examine these claims based on the spontaneous mixing behavior of bilingual speakers situated in the context of their bilingual speech community. In so doing, we test whether phonetic integration is a necessary, or even useful, diagnostic of loanword status.

Although grammar-external considerations, like the role of community conventions, or speakers' bilingual proficiency or generation, are occasionally invoked as explanations of refractory loanword adaptation strategies (e.g. Bullock & Toribio 2009, Chang 2008, Crawford 2007, Haugen 1950, Heffernan 2007, Kang 2010, Lev-Ari & Peperkamp 2014, Paradis & LaCharité 2008, Peperkamp et al. 2008, Yip 2006), the vast majority of this work has been carried out with no reference to the process of borrowing per se. From the perspective of the ground, several observations can already be made with respect to the above claims. Perhaps the most obvious is that the language mixing that characterizes bilingual communities includes not only borrowing, but a variety of other strategies as well. The cumulative results of years of research on the grammatical behavior of these mixed elements have established that only some (i.e. those that have been BORROWED) undergo integration at the morphosyntactic level; others (e.g. multi-word code-switches (CSs)) resist it. Confirming whether this dichotomy also holds at the phonetic or phonological level (as is generally assumed) in the first instance requires identifying and distinguishing these different manifestations of language contact.

---

[1] See Kang 2011 for a summary of the different strategies proposed.

Once it has been ascertained that the material under investigation has in fact been borrowed (no simple task, as it necessitates in-depth knowledge of community norms in addition to careful linguistic analysis), at least two types must be recognized: attested loanwords (ATT LWDS) and nonce borrowings (NBs). Only the latter represent spontaneous, on-line borrowing per se; it follows that this must be the privileged (and logical) locus for any study of loanword integration. And since NBs are the sole province of bilinguals, they may display very different patterns from ATT LWDS, for which no knowledge of L_D is required. This important distinction between ATT LWDS and NBs has occasionally been noted in the loanword-phonology literature (e.g. Kang 2011, Paradis & LaCharité 1997, 2008, Peperkamp 2005), but under the (untested) assumption that they behave 'similarly', they have never to our knowledge been operationally distinguished. Such failure to differentiate disparate instantiations of language mixing conflates words that are (presumably) borrowed with others that are unquestionably not (i.e. single-word CSs). This contributes to the lack of consensus surrounding the mechanics of loanword adaptation.

At least part of the problem resides in the data that form the basis of much of the aforementioned work. They derive, almost in their entirety, from two main sources: dictionary attestations or other written material (Crawford 2007, Heffernan 2007, Kang 2010, Kenstowicz & Suchato 2006, Paradis & LaCharité 2008, Rose & Demuth 2006), with all of the attendant inferences that these require about the initial processes giving rise to the eventually attested forms, and laboratory experiments (Lev-Ari & Peperkamp 2014, Peperkamp et al. 2008, Vendelin & Peperkamp 2006). Many of the latter involve foreign phone PERCEPTION; the few that target production at all obtain it through forced-elicitation procedures (Bullock 2009). It is unclear what relation either of these types of material bears to the actual process of bilingual borrowing, if any. Both inevitably abstract away from much of the variability inherent in actual production, which, as we will see in what follows, turns out to be the defining characteristic of phonetic adaptation.

A final aspect that has been almost entirely overlooked—astonishingly, for the study of phonetic realizations—is the individual factor.[2] Granted, bilingual proficiency of both community and speaker is occasionally invoked (often with reference to Haugen's (1950) three stages) and, in at least one case (Paradis & LaCharité 2008), specifically incorporated as a parameter of the investigation. But the uncritical assumption throughout is that generation or time period or intensity of contact can be straightforwardly linked with a potential borrower's bilingual ability. Close inspection of the makeup of bilingual communities confirms that this is not the case. Such communities are typically constituted of individuals with a wide range of proficiencies in the other language, ranging from passive knowledge to balanced bilingualism. Crucially, moreover, there may be no correlation between such abilities and preferred strategy (or even propensity) for integration. This is because individuals manifest a wide range of 'talents' with respect to control of L_D, as well as personal predilections as to how to handle incorporations from it. Nowhere is this more evident than in the phonetic domain. Examples abound of speakers who display native mastery of the morphology, syntax, and pragmatics of a second or target language (L_D in current terminology), while failing to produce some or even all of its phonetic inventory. If the resulting 'foreign accent' (i.e. L_R realization of L_D material) happens to surface while borrowing an L_D word, it will inevitably be identified by the an-

---

[2] Indeed, the rule- or constraint-based analyses of the loanword-phonology literature virtually demand a degree of abstraction away from individual data points.

alyst as an instance of integration. But for many speakers (or possibly all speakers under some circumstances), such INVOLUNTARY integration also occurs in (unmixed) L_D. In these cases, the only way to ascertain whether we are in fact dealing with the act of VOLUNTARY integration is by explicit comparison with the speakers' treatment of like items in both their L_R and their other uses of L_D (i.e. monolingually and in other types of language mixing).

Marshaling all of the above considerations, in this article we examine the spontaneous phonetic integration of other-language material in a principled sample of bilingual members of a well-defined bilingual speech community. Integration is understood, here as in our previous work, as the assignment of L_R characteristics to an L_D item. Although not entirely coterminous with loanword adaptation, an understanding of phonetic integration can provide crucial input to the debate over whether and how loanwords are adapted. In this endeavor, we depart both methodologically and theoretically from the existing literature. Participants were specifically selected on the basis of their relative propensity to spontaneously (nonce-)borrow L_D-origin words into their native French as well as to code-switch from French into L_D. All also make copious use of L_D-origin dictionary-attested loanwords. The linguistic analysis centers around arguably the most conspicuous phonetic differences between English and French: the acoustic realizations of voice onset time in the voiceless plosives P/T/K, and the segments represented orthographically as TH, H, and R. Their treatment in on-line borrowing from English into French (NBs) is contextualized with respect to three benchmarks: (unmixed) French, CS to English, and lexical retrieval of English material previously borrowed into French (ATT LWDS). This allows us to compare their behavior in borrowed vis-à-vis nonborrowed words. Contamination of results with involuntary integration resulting from inability to achieve L_D targets is minimized, as only participants capable of producing the diagnostic elements in both L_R and L_D form were retained. This ensures that the 'integration' we measure in fact arises from choice rather than need. Individual (e.g. age, bilingual ability) and group (e.g. neighborhood) characteristics, which have been shown elsewhere to correlate with linguistic productions, are also analyzed.

This will ultimately enable us to ascertain the utility of phonetic realization as DIAGNOSTIC of loanword status. We address the specific question, which has plagued the field for decades (Bullock 2009), of whether it works in tandem with morphosyntactic integration to determine the status of other-language material. In particular, do bilinguals marshal phonetics consistently to distinguish their NBs from other types of language mixing? And are analysts, by extension, justified in so doing? The answers will contribute to dispelling a good deal of the indeterminacy surrounding current theories of code-switching and borrowing, in addition to providing crucial on-line production data that can inform loanword adaptation theories.

The remainder of this article is organized as follows: we first describe our data and methods, paying special attention to what most distinguishes them from previous work: the focus on the borrowing process as it occurs spontaneously in the bilingual community, the distinction among different types of language mixing and the targeting of individuals known to be capable of producing L_D realizations, and segments whose alternate realizations demarcate L_R from L_D. Section 3 focuses on the status of phonetic integration as a diagnostic of language membership by comparing speakers' treatment of NBs with benchmark categories (ATT LWDS and CSs). Given the finding of pervasive variability, §4 tackles the question of which factors, linguistic and extralinguistic, promote integration. We present our global conclusions in §5.

2. DATA AND METHODS. We noted above that most of the data forming the basis for existing theories of loanword adaptation were obtained from dictionary attestations, labo-

ratory experiments, or other sources whose relation to the actual process of loanword integration is unclear. In contrast, this study relies on the spontaneous mixing of French and English recorded in a massive corpus of the bilingual speech produced by a principled sample of individuals in the context of their bilingual community. A detailed description of the OTTAWA-HULL FRENCH CORPUS, collected in the national capital region of Canada, is given in Poplack 1989. The linguistic details of participants' code-switching to and borrowing from English (Poplack 1985, 1988a,b, Poplack & Dion 2012, Poplack et al. 1988) as well as the structure of their native French (e.g. Dion 2003, Elsig 2009, Kabano 2004, Lealess 2014, Leroux & Jarmasz 2006, Poplack 2015, Poplack & St-Amand 2007, Poplack & Turpin 2009, Poplack et al. 2012, Willis 2000) have already been thoroughly documented, and community norms with respect to all of them are well understood. The size of the corpus makes it possible to locate reasonable quantities of the rare elements of interest, which may be sparsely (or not even) attested in smaller corpora. These facts make this an ideal source for the investigation of phonetic integration. In this section we describe the data we obtained and the analytical methods we employed.

2.1. SAMPLING CONSIDERATIONS. In the bilingual speech community, the conduit through which loanword integration takes place is the individual, whose personal aptitudes and proficiencies determine in the first instance their 'command' of the other language. Most pertinent to the concerns of this study is the (IN)ABILITY to faithfully reproduce other-language sounds. Failure to do so may be manifested as a 'foreign accent', typically resulting from inexact approximations of L_D forms, or outright replacement by L_R forms. L_D forms produced in L_R coincide with what we have defined as 'integration', raising the problem, rarely explicitly acknowledged (with the notable exception of Bullock's (2009) cautionary remarks), that many APPARENT integrations may simply stem from a speaker's inability to achieve target L_D-like pronunciations. Thus, any study of the phonetics of integration must painstakingly distinguish such involuntary integrations from those resulting from actual speaker choice.

Independent of speaker abilities, propensity to integrate may depend on the status of the mixed element itself. For example, once a borrowed word is no longer recognized as foreign in origin (i.e. has assumed the status of bona fide loanword), its original phonetic form will be targeted less, if at all. In this context, the number of apparent integrations will increase concomitantly, since many—if not most—L_R speakers will be pronouncing the L_D-origin word like any other L_R word, if only because they are unaware of its L_D origin (cf. 'English' *terrace*, *beef*, *jury*, among countless others). By contrast, CSs into another language are expected to display only the phonetic features of that language (i.e. no integration at all). Here too we want to distinguish integrations resulting from the unavailability or inappropriateness of the original L_D form from the products of speaker choice.

Since these limiting factors are as operative when a bilingual is speaking (unmixed) L_D as when s/he is mixing it with L_R, apparent integrations of borrowed material must be explicitly contextualized. This requires access to on-line (i.e. NONCE) borrowing and the machinery to distinguish it from lexical retrieval of previously borrowed words, on the one hand, and code-switching (by definition involving recourse to L_D grammar), on the other. Accordingly, here we compare realizations of phonetic diagnostics produced by the same individuals under four conditions: while nonce-borrowing from English, code-switching to English, drawing from their stock of attested English-origin loanwords, and speaking unmixed French (their L_R). These comparison points are schematized in Figure 1. The above considerations are occasionally invoked as caveats in the study of borrowing, but have not yet been systematically controlled for. Here they play a key role in the study design, as described below.
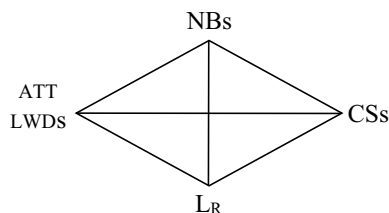
Figure 1. Contextualizing language-mixing types.

2.2. Speakers. From the random sample of 120 speakers constituting the Ottawa-Hull French Corpus, we first identified those who engaged in the most nonce-borrowing and (multiword) code-switching (as defined in §2.3). We further privileged speakers with demonstrable ability to produce the diagnostic segments (detailed in §2.2) in both L_D and L_R form. This was established through both auditory and acoustic analysis of their recorded speech. Speakers were deemed capable of producing a given segment in L_D form if their data contained at least one token realized as such in the words sampled for this study or, failing that, in any other English words produced elsewhere in their recording.[3] This ensures that integration was in fact an option for them and enables us to ascertain the conditions under which they avail themselves of it.

The resulting subsample is made up of the twenty-five individuals with above-average (vis-à-vis the entire sample) quantities of NBs (here, 5+) and CSs (20+).[4] All of them also made use of established loanwords. Note that these are not necessarily the most proficient bilinguals of the Ottawa-Hull sample, nor those with any particular phonetic talent. They are simply those who (i) made the most copious use of English-origin words distributed across all three mixing types and (ii) demonstrated the capacity to produce the diagnostic segments at least once in L_D and (of course) L_R. These criteria for participant selection are dictated by the acknowledgment that even native-like proficiency in L_D grammar is no guarantee of the same in phonetic production. Nor is proficiency necessarily correlated with quantity of language mixing. Although all sectors of the population are represented, most members of the subsample turn out to be young, highly bilingual residents of two English-dominant neighborhoods (see §2.5)—a predictable enough disproportion given our sampling criteria. For all of these individuals, phonetic integration is a matter of choice.

2.3. Mixing types. From their recorded conversations, we systematically extracted the English-origin material, retaining three categories for analysis, operationally defined as follows:

- Nonce borrowing: unattested lone English-origin item, uttered a single time by a single speaker. For each English-origin word so designated, attestation status was verified by ascertaining that it was not listed in any of six popular dictionaries of

---

[3] Practical reasons arising from our focus on spontaneous mixing behavior in social context (rather than experimental design) dictated this less-than-ideal approach. On the one hand, since fewer than a quarter of the 120 original sample members satisfied the inclusion criteria for this study, requiring categorical or sustained L_D production of even these six diagnostic segments would have drastically reduced the participant pool. On the other hand, the criteria adopted can be expected to enhance the number of French realizations, resulting in increased integration scores.

[4] The total number of tokens for these speakers alone makes up over half of the data on these phenomena for the entire corpus (286/520 NBs, 1,170/2,341 CSs).

Quebec and European French published over the last century[5] and that it had not appeared in comparable corpora representing an earlier stage (nineteenth century; Poplack & St-Amand 2007) of Quebec French. Borrowing status was determined by the independent finding (Poplack & Dion 2012, Poplack et al. 1988) that the forms were consistent with the morphology and syntax of L_R. Note that this is a much more stringent definition of NB than in previous work, where the criterion was (technical) nonce status in the corpus. The examples in 1 illustrate NBs.

- ATTESTED LOANWORD: lone English-origin item listed, in the same grammatical category and meaning, in a comprehensive dictionary of French contemporaneous with the collection of the Ottawa-Hull French Corpus (*Petit Robert*; Robert 1977). To ensure that items classified as attested were in addition (relatively) ESTAB-LISHED in the community, we further required them to have been used by at least five members of the full Ottawa-Hull sample, where data permitted (2).
- CODE-SWITCH: a MULTIWORD English utterance consisting of two or more adjacent words (excluding compounds, as per the *Canadian Oxford dictionary* (Barber 2001)), as in 3.

Realizations of the diagnostic segments described in §2.4 are compared and contrasted in each of these conditions, and contextualized with respect to each speaker's unmixed BENCHMARK FRENCH.[6] As argued elsewhere (e.g. Budzhak-Jones & Poplack 1997, Poplack 2018, Poplack & Meechan 1998b), because spoken L_R may differ in unpredictable ways from what is conventionally construed as 'standard' L_R, and because there is no reason to expect a bilingual to integrate an L_D-origin form into a variety of L_R that s/he does not speak, it is imperative that the benchmark for comparison be the speaker's own variety, as exemplified in 4.

(1) Nonce borrowings
   a. Bien moi je m'étais *booké* pour cinq z enfants quand je me suis marié.
      (105.1983)[7]
      'Well I had booked myself for five kids when I got married.'
   b. Si les motards seraient partis ça serait un vrai petit *haven* …   (051.815)
      'If the bikers would be gone it would be a real little haven.'
(2) Attested loanwords
   a. Il nous restait dix cennes pour une boîte de *popcorn*.   (065.695)
      'We'd have ten cents left over for a box of popcorn.'
   b. J'aime ça, mais je prends un *break* une fois de temps en temps.   (039.734)
      'I like it, but I take a break every now and then.'
   c. On est allés à un *bar* là.   (048.600)
      'We went to a bar.'
   d. Puis là j'ai venu à boutte de travailler dans magasin sur le *cash*.   (053.710)
      'And then I managed to work in the store on the cash.'

---

[5] *Petit Robert* (Robert 1977, 1986, 2004), *Multi* (de Villiers 1997), *Dictionnaire historique du français québécois* (Poirier 1998), *Le parler populaire des canadiens français* (Dionne 1909), *Dictionnaire québécois français* (Meney 1999), and *Trésor de la langue française du Québec* (Centre national de la recherche scientifique 2004).

[6] Benchmark French is here defined as the first ten lines of unmixed French, culled from an arbitrary point (line 500) of the corpus concordance, surrounded by a 'buffer zone' of an additional five lines of unmixed French.

[7] Examples, which contain many vernacular features characteristic of spontaneous spoken French, like the nonstandard liaison in 1, are reproduced verbatim from audio recordings of the Ottawa-Hull French Corpus (Poplack 1989). Codes in parentheses identify the speaker number and line number of the utterance. English-origin material is given in italics.

(3) Code-switches
    a. Si tu es pas capable de les battre, *join them*.   (048.2730)
        'If you can't beat them, join them.'
    b. C'est trop facile. *You can get anything you want*. Trop, trop facile.
                                                      (015.318)
        'It's too easy. You can get anything you want. Too, too easy.'
    c. On entendait parler de Vanier, c'était—*it was rough*.   (015.148)
        'We would hear about Vanier, it was—it was rough.'
(4) Benchmark French
    Mais nous-autres on achetait des cordes de bois toutes coupées, tout prêt, hein? On mettait ça dans cave.   (048.596)
      'But us, we'd buy cords of wood, all chopped, all ready, eh? We'd put that in the basement.'

2.4. DIAGNOSTICS. Assessing integration at any level depends on the successful identification of CONFLICT SITES (Poplack & Meechan 1998b): elements that can unambiguously be characterized as belonging to L_D or L_R. Although the French-English language pair offers a reasonable number of phonetic conflict sites (see Poplack et al. 1988 for detail), most turned out to be unsuited to the type of quantitative empirical study we undertake here. Many apparent conflicts have large regions of overlapping (i.e. nondiagnostic) values that preclude straightforward classification as French or English. This is particularly true of the vowel system and is also the case for VOICE ONSET TIME (VOT) in voiced plosives.[8] Other conflicts (lexical stress, intonation, diphthongization) are similarly limited by the difficulty in distinguishing unambiguously English and French realizations through phonetic means alone. Additional potential diagnostics were discounted due to low frequency resulting from disproportionate distributions of segments, whether in the languages overall or in the NBs and CSs remaining once the sampling criteria were satisfied. We ultimately retained four diagnostics that capture measurable and conspicuous (arguably the most conspicuous) differences between French and English. These are the phonetic realizations of the segments represented orthographically as TH, H, and R, and the (acoustic) measurement of VOT in the voiceless plosives P/T/K. English-origin words containing these conflicts are illustrated in 5.

(5) a. J'aime pas des **th***ieves*. J'ai jamais été voleuse.   (041.807)[9]
        'I don't like thieves. I've never been a thief.'
    b. Non, mais ils te gardent sur **h***old* pas mal longtemps.   (037.1152)
        'No, but they keep you on hold for a long time.'
    c. Ah moi j'écoute pas ça, c'est assez *dep***r***essing*.   (037.1910)
        'Oh, I don't listen to that. It's too depressing.'
    d. Puis Ottawa commence à être bien **p***olluted*.   (041.833)
        'And Ottawa is starting to get real polluted.'

From the recorded discourse of each speaker in the subsample, we extracted every token of these diagnostics appearing in each of the three language-mixing types (NBs, ATT LWDs, and CSs to English), as well as in the control, benchmark French. Based on

---

[8] Caramazza and Yeni-Komshian (1974), for example, show the English values of B/D/G to be completely subsumed within those of French. This lack of English-specific values invalidates voiced plosives as a 'conflict site'.

[9] Segments under investigation are indicated in boldface, except where excluded from ensuing calculations due to poor audio quality (as the K in **p***op*c*orn* in 9b) or occurrence in coincidence sites (where their realization is not diagnostic of language membership).

a binary coding system, each was identified as INTEGRATED (realized in French) or UN-INTEGRATED (realized in English). For example, since interdental fricatives (5a) do not figure in the French consonant inventory, any segment realized with a combination of frication noise and interdental articulation was coded as unintegrated, while voiced or voiceless plosive realizations of English TH were considered integrated.[10] This is a QUALITATIVE conflict site. Other conflicts, though theoretically also qualitative, are in actuality quantitative, insofar as they share some like realizations. For example, fricative H and approximant R (both said to be nonexistent in French) do occur sporadically in native words (e.g. *dehors* [h], *docteur* [ɹ]), albeit only in a small set of lexical items and/or in certain (usually nonstandard) dialects. Although this does not apply to the members of our subsample, we nonetheless coded fricative realizations of H and approximant realizations of R in English-origin words as unintegrated. A null realization of H and a trilled or uvular fricative realization of R were considered integrated. Likewise, although there is an acoustic zone in which French and English values for VOT in P/T/K overlap (making this a coincidence site that is silent as to loanword or codeswitch status), other values were found to be exclusive to one or the other. Based on the information in Caramazza and Yeni-Komshian's (1974, figure 4) study of VOT in bilingual and monolingual Canadian French and English, we delimited, for each segment, the values that could be classified as French-specific (i.e. integrated) (P: < 35 ms, T: < 30 ms, K: < 40 ms) vs. English-specific (P: > 40 ms, T > 40 ms, K > 55 ms).[11]

Given the naturalistic circumstances under which the data were recorded, many otherwise legal tokens turned out to be indeterminate, if not altogether inaudible. Others were 'ambiguous' in the sense of being nondiagnostic (e.g. some occurrences of P/T/K that are never aspirated in L_D (such as following [s] in *toothpaste* and *high school* in 6b and 7 below) are therefore silent as to mixing type). These were excluded from further calculations. But since no other 'intermediate' realizations (in the sense of being neither French nor English) were attested, it was possible to straightforwardly categorize the remaining tokens as integrated or unintegrated, using the criteria outlined above. TH, H, and R could be coded impressionistically, since only cases that were clearly identifiable using auditory measures were retained. Differences in VOT, which are less readily discernable by auditory measures alone, required spectrographic measurement (using Praat; Boersma & Weenink 2016) of the duration in milliseconds between the plosive burst and the onset of vocal-fold vibration.[12]

The above measures of phonetic integration thus operationalize two manifestations of loanword adaptation. In one, an illicit L_D phoneme is substituted for a licit L_R counterpart (this is the case of the French [t] realization of English [θ], or [d] realization of [ð]); in the second, an illicit allophonic realization of a licit L_D phoneme is substituted

[10] No alveolar fricative [s, z] realizations of interdental fricatives, often reported in European French dialects, were found in this context.

[11] Since published VOT values reflect stressed syllables only, it could be argued that the study should be restricted to this environment. But speakers also have the choice to integrate at this level, and we cannot predict whether they will choose to apply English or French stress-assignment rules. Furthermore, 74% of the syllables that would be unstressed in English did in fact occur in stress-bearing positions for French. We therefore did not exclude a priori contexts that would be unstressed in English. As it turns out, stress is not an issue, since few tokens (12%; $N = 37$) occurred in positions that would be unstressed in both languages, and those were no less likely to be realized with English VOT values than their counterparts in English stress-bearing syllables (24% vs. 25%).

[12] In contexts where /t/-flapping is mandatory in English, flapped realizations were considered unintegrated, and other realizations integrated.

for one that is licit in L_R (as is the case for aspirated allophones of the phonemes P, T, and K when produced as [p], [t], and [k] in French).[13]

These procedures yielded a total of 1,354 tokens, exhausting the pertinent data available for the subsample. These are nonetheless quite unevenly distributed according to both diagnostic and mixing type, as depicted in Table 1.

TYPE	P/T/K	R	TH	H	TOTAL	AVG/INDIVIDUAL
Nonce borrowings	75	129	8	13	*225*	9
Code-switches	88	104	90	79	*361*	14
Attested loanwords	147	152	37	84	*420*	17
French	190	124	6	28	*348*	14
TOTAL	*500*	*509*	*141*	*204*	*1,354*	*54*

TABLE 1. Distribution of data by language-mixing type and phonetic diagnostic.[14]

Considering the dearth of studies of phonetic integration of spontaneously borrowed material in the loanword literature, it is worth clarifying why this should be. First, even in a multimillion-word data set such as the Ottawa-Hull French Corpus, lone other-language items of all frequencies represent less than 1% of the data (Poplack et al. 1988:49). NBs, which by definition occurred only once, necessarily reduce the pertinent data pool further. And multiword CSs, here as elsewhere, are even less frequent than single-word incorporations. To make matters worse, the diagnostics themselves occur at disproportionate rates in the two languages; some are extremely infrequent, especially in the relevant linguistic environments. Even where segments recur often in L_D, there is no guarantee that speakers will choose to borrow or switch specific words containing them. This explains the paucity of data (already predicted by Bullock (2009)) in some of the cells in Table 1, despite the fact that the pertinent tokens were exhaustively extracted from the recordings of each individual. In the context of this field, however, even this relatively limited amount of data is substantial (if not vast), given that we are dealing with natural production. Indeed, the analyses we report here would have been inconceivable in the absence of a bilingual resource the size of the Ottawa-Hull French Corpus, which not only is rich in other-language material but also offers critical information regarding its frequency of occurrence and dispersion across the community. On the plus side, a single lexical item may contain several diagnostic segments simultaneously (e.g. *toothpaste*). Every usable occurrence of the relevant segments was taken into account, resulting in a respectable average of fifty-four analyzable tokens per speaker.

2.5. EXTRALINGUISTIC FACTORS. Three potentially explanatory extralinguistic factors were also examined for each speaker: age, level of bilingual proficiency, and intensity of contact/official status of French at the local (i.e. neighborhood) level. These are il-

---

[13] On the question of whether these proceed via category-wise matching of phonemes (e.g. Paradis & LaCharité 1997, 2008) or through (mis)perception (Peperkamp 2005, Peperkamp & Dupoux 2003), we remain agnostic. Our study was not designed to address this issue.

[14] Note that words containing H and TH were so scarce among ATT LWDs that it was necessary to relax the attestation criterion (§2.3) somewhat. For H, we retained dictionary-unattested types provided they had recurred across five speakers or more in the Ottawa-Hull French Corpus; to be able to perform any comparison at all on TH, we were obliged to consider as 'established' any recurrent word. This should have no effect on the results reported here, since subsequent calculations (not shown) revealed that integration rates for H were unaffected by attestation. The fact that TH was no less integrated than segments in more widely diffused words (§3.4) suggests that results were not compromised here either. In addition, due to the difficulty of locating individuals fulfilling such stringent inclusion criteria even in a sample of 120 speakers, we were obliged to retain three (040, 044, 062) who, despite complying with all other requirements, produced no interdental fricative realizations of TH.

lustrated in Table 2. Such factors, described in detail elsewhere (Poplack 1989, 2018), represent complementary measures of contact; they have already been found to affect borrowing rates and type in this community (Poplack 1985, 1988b, Poplack et al. 1988). In §4 we investigate whether they have an influence on phonetic integration as well. All other things being equal, the expectation is that L_D realizations will be favored among individuals experiencing the most contact with that language, situated to the right in Table 2: the young, the most proficient, as measured by a CUMULATIVE ENGLISH PROFICIENCY INDEX (CEPI;[15] Poplack 1989), and those residing in Ottawa (especially in the higher-contact neighborhoods of Basse Ville and West End), where English is the official and majority language.

FACTOR GROUP	RELATIVE INTENSITY OF CONTACT WITH ENGLISH		
	← lower		higher →
Speaker age	OLDER (57–73)	INTERMEDIATE (39–54)	YOUNGER (17–32)
Level of bilingual proficiency	LOW (CEPI levels 1 and 2)		HIGH (CEPI levels 3 and 4)
Status of French/density of contact in the neighborhood	MAJORITY (Quebec) Mont Bleu Vieux Hull		MINORITY (Ottawa, Ontario) Vanier Basse Ville West End

TABLE 2. Extralinguistic factors instantiating intensity of contact with English.

Summarizing, the design of this study simultaneously controls for speaker, language-mixing type, and segment. We distinguish the individual from the group, involuntary integration from voluntary, and we situate the treatment of L_D material with comparative benchmark data. These methodological innovations will enable us to uncover the facts of phonetic integration in bilingual usage and to clarify whether speakers employ it strategically to differentiate borrowed material from code-switching.

3. ASSESSING PHONETIC REALIZATION AS A DIAGNOSTIC OF LANGUAGE-MIXING STATUS.
3.1. EXPECTATIONS. Before plunging into the results of this study of integration at the phonetic level, it will be instructive to review the treatment accorded L_D material at other linguistic levels. Decades of research, in this and other bilingual communities, have shown that words that have been borrowed, be they NBs or long-attested loanwords, show an overwhelming tendency to assume the morphology and syntax of L_R, to the extent of reflecting its inherent variability and associated conditioning (Poplack 2018). This is not a tendency, but a robust finding, based on large amounts of data from many language pairs (e.g. Spanish/English (Torres Cacoullos & Aaron 2003), Ukrainian/English (Budzhak-Jones 1998), Igbo/English (Eze 1997), Persian/English (Ghafar Samar & Meechan 1998), Tamil/English (Sankoff et al. 1990), Turkish/English (Adalar & Taglia-

---

[15] Bilingual proficiency is notoriously difficult to measure. Here we use it as an umbrella term for the relative importance of English in a native francophone's repertoire. To assess this we calculated, for each speaker, a CEPI based on self-assessments, largely gleaned through content analysis, on a number of complementary measures: age and means of acquisition, current frequency of use, language chosen with critical interlocutors, and reported and observed production. Each response was weighted according to its contribution to the overall strength of English in the speaker's repertoire. Thus, an individual who learned English spontaneously in early childhood would have scored higher on the acquisition measure than one who acquired it at a later age or in a formal institution. Likewise, choice of English in the home setting counted more heavily than in the workplace. Based on these calculations, each sample member was assigned an average cumulative score between 0 (no knowledge/use of English) and 1 (high proficiency). The threshold for HIGH was arbitrarily set at 0.46 to ensure that statistical comparisons would be meaningful. CEPI scores are by no means to be interpreted as absolute; they simply capture, for each speaker, the relative importance of English.

monte 1998), Finnish/English (Wheeler 1987), German/English (Nault 2002), Gulf Arabic/English (Mustafawi 2002), Tunisian Arabic/French (Poplack et al. 2015), English/Acadian French (Turpin 1995), and Wolof/French and Fongbe/French (Meechan & Poplack 1995, Poplack & Meechan 1995)). Moreover, such grammatical integration occurs at the first mention of the word, that is, at the NB stage (Poplack & Dion 2012). Unambiguous CSs to L_D, by contrast, retain the morphology and syntax of L_D. This means that spur-of-the-moment borrowings need not undergo a lengthy adaptation period before achieving the linguistic appearance of bona fide loanwords.[16]

In this context, for phonetic realization to be deemed a useful diagnostic of language-mixing type, our expectations are the following:

A. Phonetic realization should be consistent with the L_D item's grammatical status: as with the morphosyntax, ATT LWDS and NBs should reflect the behavior of the L_R benchmark, and code-switched items should mirror that of L_D, including any variability either may display. Here, because the segments under study were selected specifically on the basis of their diagnosticity (§2.4), this means that in NBs and ATT LWDS, TH, H, R, and VOT in voiceless P/T/K would be produced in French (as [t]/[d], Ø, [χ], and [p]/[t]/[k], respectively), while the same segments occurring in CSs to English would display English realizations ([θ]/[ð], [h], [ɹ], and [pʰ]/[tʰ]/[kʰ]).

B. The considerations in A should equally apply to all speakers (given the criteria for inclusion in this study) and to each of the three mixing contexts. Individuals should realize their NBs and ATT LWDS in L_R form, and CSs in L_D form, again paralleling the relevant benchmarks, including any variability they may feature.

C. The considerations in A should apply equally across linguistic environments; that is, propensity to integrate should not differ according to segment.[17]

3.2. PHONETIC REALIZATION OF THE RECIPIENT-LANGUAGE BENCHMARK. As noted above, to conclude that phonetic realization is a reliable measure of mixing type, we require that NBs display a phonetic-integration rate parallel to that of the L_R benchmark (expectation A). Bearing in mind that the relevant characteristics of L_R may differ substantially (and unpredictably) from inferences based on the standard variety of that language (Budzhak-Jones & Poplack 1997, Sankoff et al. 1990), we first investigate how the diagnostics are realized in these speakers' benchmark French. We in fact find them to be almost (though not entirely) categorically (93%) produced in the latter (Table 3).[18]

SEGMENT	%	*N*
R	100	124/124
TH	100	6/6
H	100	28/28
P/T/K	88	167/190
TOTAL	*93*	*325/348*

TABLE 3. Realization of diagnostic segments as French in L_R benchmark data.

[16] Nor is this even a realistic scenario, despite its widespread acceptance; the same study shows that the vast majority of NBs are ephemeral, meaning that they do not even persist long enough to participate in any gradual adaptation process.

[17] Note that expectation C applies only to integration into L_R. A variety of factors (e.g. motoric difficulty, psycholinguistic considerations, etc.) may of course affect propensity to produce segments in L_D (English).

[18] Only P/T/K deviate slightly, showing some variability between English and French realizations, reflecting ongoing contact-induced change in the community.

3.3. PHONETIC REALIZATION OF NONCE BORROWINGS. Speakers who realize an L_R phonetic element categorically in a particular form should do the same for counterparts contained in words borrowed into L_R. As a corollary, we further expect that any variability in L_R should also be mirrored in NBs and ATT LWDS. Given these expectations, NBs should virtually always be realized in French (at 93%). In striking contrast, however, the NBs studied here are integrated barely a quarter of the time (24%; $N = 59/225$) in the aggregate. This is a dramatic departure from both expectation A and the near-categorical integration observed of the same items on the morphological and syntactic levels (§3.1). Such robust variability in and of itself detracts from the utility of phonetic realization as an indicator of loanword or code-switched status: indeed, reliance on phonetic form as a sole diagnostic would lead us to misclassify most NBs.

INTEGRATION ACROSS INDIVIDUALS. Recall, however, that this amalgamated rate conflates values from twenty-five speakers with varying extralinguistic characteristics. If variability in phonetic integration of NBs were due to amalgamating speakers for whom phonetic form is in fact diagnostic (as per expectation A) with others for whom it is not, its ability to serve as an indicator of mixing type could be at least partially salvaged, especially if the individuals in question belonged to identifiable extralinguistic cohorts. We review overall integration rates across speakers as a preliminary gauge of this possibility.

Not one speaker complies with expectation A. Three quarters of the sample display variability (at rates ranging from 11% (speaker 013) to 78% (048)), while the 24% who do treat their NBs categorically (speakers 007, 015, 056, 062, 066, and 082) opt not to integrate (realizing their NBs in L_D)—the diametric opposite of what would be required if phonetic realization were diagnostic. The extent of intra- and interspeaker variability observed here (contra expectations A and B), coupled with the generally low integration rates across speakers (29% average, 30% median), bolsters the suggestion that the unified community-wide strategy of integrating NBs at the morphosyntactic level does not apply at the phonetic level. Indeed, inclusion of the ninety-five community members of the wider sample who did not meet the criteria for this study would only further extend this range.

Since distributional analyses of the extralinguistic factors of age, bilingual proficiency, and intensity of contact at the local level considered independently neither corresponded to the expectations summarized in connection with Table 2 nor yielded another straightforward explanation,[19] in §4 we adopt a mixed-effects model to revisit

---

[19] As an example, consider the social characteristics of the most extreme (non)integrators in the sample. The social profiles in (i) show that not only is the social makeup of each class heterogeneous, but speakers who do share a social profile (056/048 and 062/040) also adopt opposing integration strategies, the first of each pair belonging to the cohort of nonintegrators, the second figuring among the highest integrators.

(i) Speaker	Integration rate	Age	Bilingual proficiency	Intensity of contact in the neighborhood and province
NONINTEGRATORS				
007	0%	54	High	Intermediate (Vanier, ON)
015	0%	27	High	Intermediate (Vanier, ON)
056	0%	57	Low	High (West End, ON)
062	0%	25	High	High (West End, ON)
066	0%	45	Low	High (West End, ON)
082	0%	70	High	Low (Vieux Hull, QC)
HIGHEST INTEGRATORS				
105	50%	64	Low	Low (Mont Bleu, QC)
040	64%	31	High	High (Basse Ville, ON)
048	78%	73	Low	High (Basse Ville, ON)

the role of the individual in conjunction with these factors. We note, however, that while the results will serve to illuminate the factors conditioning the variability we have documented, they do not mitigate the finding that phonetic realization is a poor indicator of the status of NBs: no speaker's realizations parallel that of the target L_R benchmark. Most deviate dramatically from it, some to the point of not integrating at all.

INTEGRATION ACROSS SEGMENTS. Examination of integration rates by segment is even more damaging to our expectations. These suggest that for many speakers with variable realizations, phonetic integration of NBs is also mediated by the segment involved, thereby contravening expectation C. Each of those considered here is realized not only variably in L_R, but also to apparently wildly differing extents: aggregate rates range from very rarely for R (9%) and H (10%) to much more often for P/T/K (50%) and TH (70%). This suggests that different segments are differentially amenable to integration, a result also reported by Lev-Ari, San Giacomo, and Peperkamp (2014) for a different contact situation and language pair. Here, independent of other considerations, R and H tend to be produced in English, while P/T/K and TH are more likely to be realized in French. Since many NBs contain more than one diagnostic segment, this extends the potential for variability even to the level of the word.

By way of illustration, consider that 33% (52/160) of the NB types analyzed in this study contain more than one diagnostic segment: forty-one have two (e.g. **h**eartbea**t**, **r**ob**b**ers), and eleven have three or more (e.g. **h**un**t**ers, **airp**ort, **p**re-**k**inde**r**garten). Rather than treat the segments constituting the word as either L_R (or L_D), the two coexist in the same word a third of the time, as in 6. The more conflicts in the word, the more inconsistent the realization (29% integration for words with two segments, 45% for those with three or more).[20] And since many L_D-origin words contain conflict sites beyond the four targeted here, considering more of them would likely have turned up further inconsistency. We return to the influence of segment in the statistical analysis of §4.

(6) a. **p**owe**r**   [p], [ɹ]   (FR, ENG)   (025.3020)
    b. **t**oo**th**paste   [tʰ], [t], [p]   (ENG, FR, FR)   (014.774)

SUMMARY: NONCE BORROWINGS. Summarizing, examination of speakers' phonetic treatment of diagnostic segments in NBs revealed that, in the aggregate, these are integrated into French less than a quarter of the time (contra expectation A)—a surprisingly low rate given the near-categorical integration of the same NBs at the morphosyntactic level (Poplack 2018, Poplack & Dion 2012). A variety of breakdowns uncovered major fluctuations across speakers (contra expectation B), segments, and even within the confines of a single word (contra expectation C). This militates against the utility of phonetic realization as a diagnostic, at least for NBs.

It could be argued that these puzzling results are in some respect attributable to the properties of NBs themselves. Recall that these by definition are neither recurrent nor widespread, having been uttered only once by a single speaker. This could theoretically have interfered with the establishment of a shared phonetic strategy (although it does not affect their near-categorical integration at the morphosyntactic level). But before speculating about the interpretation of these results, it will be useful to ascertain whether the benchmark mixing types are themselves behaving according to expectation in this domain. Accordingly, in the next section, we replicate our analysis on ATT LWDS and CSs.

---

[20] Of course, the likelihood of inconsistency necessarily increases as a function of the opportunities for it, but this fact only compounds the unreliability of phonetic realization as a metric.

3.4. PHONETIC REALIZATION OF BENCHMARK L$_D$-ORIGIN ITEMS.

ATTESTED LOANWORDS. ATT LWDS, by all accounts, are bona fide members of the French lexicon, despite their English etymological origin. As explained in §2.3, virtually all words here qualified as ATT LWDS not only are dictionary-attested, but have also achieved a certain level of diffusion across the community. Established loanwords should assume the relevant linguistic characteristics of their L$_R$ counterparts (expectation A), and it has been shown for these selfsame words that this is in fact the case with respect to their morphosyntax (Poplack & Dion 2012). Surprisingly, however, with an overall integration rate of only 56% ($N = 234/420$), they display no particular propensity toward integration at the phonetic level (contra expectation A). This means that their phonetic form misrepresents their documented status nearly half of the time.

To complicate matters, as with NBs, this aggregate rate obscures substantial interspeaker differences, contravening expectation B. Individual rates range from 0% to 100% (average 56%, median 58%), with a full 88% displaying variability. Only two of the twenty-five speakers examined treat ATT LWDS as per general expectation, by integrating them consistently into French. One even realizes all of them as English.[21] We also observe the differential associations of integration with segments already noted for NBs (contra expectation C), including the tendency of TH and P/T/K to be integrated far more (74% and 76% respectively) than H and R (31% and 44%).

We had suggested earlier that the seemingly erratic behavior of NBs could conceivably be a consequence of their status as nonce words, making the finding that it is echoed in ATT LWDS—which by definition are recurrent—all the more perplexing. We can capitalize on the fact that ATT LWDS feature multiple iterations to clarify whether consistent phonetic treatment at least obtains at the level of the word. Examination of the ninety instances where the same ATT LWD was uttered more than once by a single speaker suggests otherwise. This is illustrated in 7, where speaker #004 produces *high school* twice with an initial fricative and twice without.

(7) **h**igh school [h] (ENG) (004.1181; 004.1562) / Ø (FR) (004.1761; 004.1762)

Such vacillation is even more apparent in words with multiple conflicts, despite our very generous interpretation of consistency, here including not only integrating (8a) or failing to integrate (8b) a segment at each iteration of the word, but also repeating the inconsistent pattern of integrating one but not the other, as in 9.

(8) a. **th**ough   [d]   (FR)   (065.504; 065.667; 065.1232)
    b. **th**ank you   [θ]   (ENG)   (013.11; 013.1173; 013.1499)
(9) a. **t**eenage**r**s   [t], [ɹ]   (FR, ENG)   (053.65; 053.222)
    b. **p**opco**r**n   [p], [ɹ]   (FR, ENG)   (025.1397; 025.1398)
    c. **th**e**r**apies   [t], [ɹ]   (FR, ENG)   (065.310; 065.313)

Even where a speaker shows allegiance to a specific phonetic rendition of a segment across iterations of the same loanword, there is no guarantee that that segment will be realized the same way in a different one. Thus, speaker #068 utters *horn* twice without integrating either segment, but integrates the two segments in *cafeteria* both times; speaker #014 consistently realizes *traffic* with French [χ] and *rock* with English [ɹ]. These examples could be multiplied. Only two speakers realize the same segments in the same way across all of their recurrent ATT LWD types, but even they do so by means of opposing strategies—one always produces the segments as English, while the other always integrates them into French. And comparison of interspeaker realizations of the same

---

[21] And this speaker is not even one of those who applied this strategy to his NBs.

ATT LWD, predictably enough, multiplies the inconsistency. Even by conservative estimates,[22] a full third of the ATT LWDs with a minimum level of diffusion (i.e. recurrence in the speech of two or more individuals) display differing interspeaker realizations.

The case of *hockey*, a loanword attested since 1889, and, at 237 tokens uttered by sixty of 120 speakers in the wider Ottawa-Hull French Corpus, the most widely diffused in our study, is perhaps most illustrative. Even this old and widespread loanword fails to enjoy uniform intraspeaker or community-wide realization, and the same is true of its two diagnostic segments (Table 4). Realizations of H and K are categorically French for some speakers (H = #s 038, 066, 105; K = 007, 004, 066, 105), always in English for others (H = 037, 014, 051; K = 056, 014), and variable for still others (H = 070, 057, 004; K = 025, 051, 057). And even those few speakers who produce the two segments consistently may treat them in contrasting fashion, integrating both (066, 055) or neither (014).

SPEAKER	H		K		OVERALL
	%	CONSISTENCY	%	CONSISTENCY	
037	0	✓	—		✓
070	50	X	—		X
038	100	✓	—		✓
056	—[a]		0	✓	✓
025	—		67	X	X
007	—		100	✓	✓
014	0	✓	0	✓	✓
051	0	✓	67	X	X
057	50	X	33	X	X
004	50	X	100	✓	X
066	100	✓	100	✓	✓
105	100	✓	100	✓	✓

TABLE 4. Consistency in integration of ATT LWD *hockey* (individuals with two or more tokens).
[a] Poor audio quality precluded measuring more than one token where indicated, meaning that consistency could not be established.

Summarizing, ATT LWDs display the same kinds of fluctuation in integration across speakers and segments as observed for NBs. This is already remarkable, given both the widespread expectation that they be fully integrated into L_R and the independent evidence that this is in fact the case on the morphosyntactic level (expectation A). Recall that as native francophones, all members of the community under study (whether drawn from the subsample or the wider corpus) are of course capable of integrating these segments (i.e. producing them in French), those with least proficiency in L_D doing so de facto. Consistent integration should therefore not be problematic to achieve, and of all mixing types, ATT LWDs would be predicted to be most amenable. These results, taken together with those for NBs, thus reinforce our contention that phonetic integration is not a reliable indicator that L_D words have been borrowed into L_R.

MULTIWORD CODE-SWITCHES TO ENGLISH. Turning now to the final benchmark, we recall that whereas ATT LWDs are construed as full-fledged French words, those contained in CSs as we have defined them are uncontroversially English. As such, they should be governed by the grammar of English, including, presumably, its phonetic requirements (expectation A). In this context, their observed rate of phonetic 'integration'

---

[22] Capping the tokens considered to a maximum of three for the eighteen ATT LWD types recurring at least twice per speaker for at least two speakers (for a total of 191 tokens). Consistency would likely further diminish if more ATT LWD types or tokens were available, and/or if the level of diffusion were greater.

into French (38%!; $N = 137/361$) is far off the expected rate of 0% (and indeed even higher than that of NBs). This is particularly puzzling in view of participants' demonstrated capacity to produce target L_D realizations.

Here as well, the amalgamated rate reflects much intra- and interspeaker variability. Rates differ drastically from one individual to another, ranging from a low of 11% to a high of 75% (38% average and median), contra expectation B. Not one speaker realizes the English words contained in CSs categorically in English—although some do so for their NBs and even their ATT LWDs.

Integration of CSs again differs across segments (contra expectation C); but instead of the bimodal distribution observed for NBs and ATT LWDs, the four studied in this context appear positioned at distinct points along a continuum. P/T/K most often misrepresent the status of the CSs in which they occur, with 67% of cases produced as in French; TH is nearly as likely to be realized in French (52%) as in English; and one third of the HS are null, as per French. Only R comes close to accurately reflecting the switch to English, nearly always (97%) being realized in that language.

In sum, CSs do not conform to (analysts') expectations of their behavior any more than NBs or ATT LWDs. The fact that they are so frequently realized in French might be understandable were we dealing with the larger community, given the wide array of proficiencies and proclivities. But these calculations are based only on the subsample of speakers defined by the ability to produce L_D forms. Admittedly, some L_D segments may be 'easier' to produce than others, and speakers may also have different thresholds beyond which they cannot sustain accurate L_D realizations. Psycholinguistic factors promoting differential realizations are undoubtedly operating as well.[23] But these considerations do not explain why the variability, both interspeaker and intersegment, that characterizes the treatment of CSs (production target L_D) parallels that obtaining in NBs and ATT LWDs (production target L_R). All of these facts are detrimental to the assumption that phonetic integration is a necessary or useful adjunct to determining the status of L_D-origin material.

4. COMPETING INFLUENCES ON PHONETIC INTEGRATION. The distributions discussed in §3 revealed extensive variability across diagnostic segments and individuals, not only for NBs, but surprisingly for benchmark ATT LWDs and CSs as well. Rather than parallel the phonetics of L_R when borrowing and of L_D when switching, as per expectation (and their observed behavior at the morphosyntactic level), the L_D-origin items studied here show no particular allegiance to either. This means that mixing type in and of itself is not determinative of phonetic realization. But we have yet to ascertain what role it plays, if any, when considered in conjunction with other factors that appeared from the foregoing distributions to affect all three types of L_D item. In this section we pool these types to carry out a logistic regression analysis with mixed-effects modeling. This will uncover which factors affect choice of form in a statistically significant way, and will determine their relative importance. While the results cannot impeach our conclusion that phonetic realization is at once too variable and too divergent from the relevant benchmarks to constitute a failsafe diagnostic of language-mixing type, they will clarify the role of mixing type and other potential linguistic and extralinguistic contributors to the observed variability.

4.1. HYPOTHESES. The major expectation underlying this research was that mixing type would be fully determinative of phonetic form (expectation A), just as it is in the

---

[23] As, for example, with the otherwise inexplicable avoidance of apical or velar (i.e. integrated) realizations of R in Canada versus the overwhelming preference for the latter in Europe.

morphosyntactic domain. Although the results of preceding sections suggest that this is not the case, we hypothesize that mixing type will emerge not only as statistically significant but also as the most important predictor. Expectation A is hypothesized to be reflected quantitatively, such that integration would be favored in ATT LWDs and NBs (HYPOTHESIS 1; H1). Expectation B was likewise countered by the varying integration rates by speaker revealed in §3; here we hypothesize that whatever interspeaker variability exists should be explicable by membership in extralinguistic cohorts defined by contact-related parameters detailed in Table 2. Specifically, we expect a greater likelihood of L_D (i.e. nonintegrated) realizations from the young, the most L_D-proficient, and those residing in high-contact neighborhoods (HYPOTHESIS 2; H2). The expectation that there would be no distinction among segments in propensity to integrate (expectation C) was likewise thwarted by the observed distributions. Because a segment effect is the most problematic for any claim that phonetic realization is diagnostic, we hypothesize that its contribution will be minimal (HYPOTHESIS 3; H3). As far as differential propensity of specific segments to integrate is concerned, we have no hypothesis at this time. As observed above, though they may be differentially difficult for speakers to produce in English, all should be equally easy to produce in their native French.

4.2. SUMMARY OF DATA. As described in §2, the data comprise 1,006 tokens[24] of L_D-origin items occurring in three mixing types produced by twenty-five speakers who are capable of both English and French realizations of the six targeted segments. For each, we distinguish neighborhood of residence, age, and relative bilingual proficiency. Phonetic realization of each instance of the diagnostic segments is coded as integrated (produced in French) or not (produced in English).[25]

4.3. MODEL SUMMARY. The number of divisions described above—the minimum dictated by our research questions—inevitably results in unbalanced and/or sparsely populated cells. For readers unaccustomed to working with spontaneous bilingual speech, it is worth stressing that there is no easy remedy for this problem: the 1,006 tokens retained exhaust all of the possibilities offered by the multimillion-word Ottawa-Hull French Corpus once the parameters of the study are respected. There are no more data where those came from. We can only speculate on how refractory to quantitative analysis the data would have been if extracted from an average-sized corpus—and how misleading the results, if based on no data at all. The scarcity and imbalance of the data points in this data set, coupled with the idiosyncratic speaker behavior described in §3, justify our choice of a mixed-effects model. Incorporating individual-level random effects in quantitative analyses of linguistic phenomena is particularly important when group-level factors, such as community, age, and bilingual proficiency, are hypothesized to be involved.

Given the presence of distributional imbalances in naturally occurring data (and in this study specifically), the challenge was to identify a model with adequate complexity to represent the full range of possible contributors to the attested variation, while ensuring that the fixed and random effects could be estimated reliably. Logistic regression models with mixed effects (Baayen et al. 2008, Bates 2010, Pinheiro & Bates 2000) were estimated using the function glmer() from the R package lme4 1.1-11 (Bates et al. 2015). Data balance may be problematic if it causes estimation not to converge; model-profile diagnostics cannot be computed in lme4 without convergence. Hence, convergence is required for models based on this data to be meaningful. All models run here

---

[24] As the statistical analysis focuses on L_D-origin items, benchmark French tokens ($N = 348$) are set aside.
[25] The full cross-tabulation appears in Appendix A1.

include an overall intercept term, corresponding to the 'input value' or 'corrected mean' of Varbrul models.[26] The dependent variable is phonetic realization; the variants are French and English form. The success value (in variationist terms, the 'application value') is the French variant. Positive parameter values are interpreted as 'favoring' integration into French.

Within linguistics, there are conflicting recommendations as to exactly how to apply the mixed-effects framework, with differences regarding the units for which random effects are required. Specifically, Barr et al. 2013:255 advocates employing a model with random intercepts and slopes for 'the maximal effects structure JUSTIFIED BY THE DESIGN' (emphasis in the original), the benefit of which is that linguistic effects observed in the aggregate across speakers are based on observations of such effects existing within individuals. Random intercepts alone do not accomplish this. In contrast, Baayen et al. 2015, Bates et al. 2015, and Matuschek et al. 2017 caution that empirical data balance should be taken into consideration in model specification, and that the 'maximal' recommendation may lead to model overspecification and increased risk of type II error. In other words, requirement of random slopes in a model where the empirical data are not sufficient to support them could increase the risk of failing to observe linguistic effects when they are actually present in the data. For our data, using lme4, none of the models with random slopes that we have considered converges, convergence failure being one diagnostic for model overspecification (Bates et al. 2015). For this and other reasons, our model specification includes only a single per-speaker random intercept and no random slopes.[27] This specification cannot address questions involving per-speaker type or segment effects, or group-level interactions with type and segment that might depend on them. However, in a Bayesian fit of a model with by-speaker random slopes for segment and mixing type using the R brms package interface to Stan (reported in Appendix A2; Bürkner 2017, Stan Development Team 2018), none of the qualitative conclusions regarding the fixed effects of interest differ from those in the model presented here.

As regards the fixed-effects model structure, we include the factor of mixing type, alongside the linguistic factor of segment and extralinguistic factors of age cohort, level of English proficiency, and neighborhood. We therefore adopt the formula in 10 as the model with the largest set of factors we considered reasonable to estimate on our data.

> (10) MODEL 1: integration ~ 1 + type + segment + neighborhood + proficiency + age + (1 | speaker)

The response variable (integration) is a binomial representing the number of realizations of the variable of interest in French versus English phonetic form; the predictor variables are the intercept (symbolized as 1), the type of item (ATT LWD, NB, or CS), the diagnostic segment (P/T/K, TH, H, or R), and the speaker's neighborhood (Mont Bleu, Vieux Hull, Vanier, Basse Ville, or West End), proficiency (higher and lower), and age (younger, midrange, and older), along with a random intercept for speaker (1 | speaker).

4.4. SIGNIFICANCE TESTING. In order to evaluate which factors in model 1 are relevant to integration in a statistically meaningful way, we use the standard method of backward elimination (Agresti 1996, Gelman & Hill 2007), in which a suite of models based on model 1 but each missing a different one of its factors is run and compared

---

[26] In the interest of consistency here and throughout §4, we use variationist terminology where possible.

[27] The recommendations of Baayen et al. 2008 and Barr et al. 2013 might suggest by-lexical-item random effects, but there are few lexical items that recur as tokens in our data set, meaning that such random effects would be difficult to estimate, and they would be relatively unlikely to dramatically affect inferences about factors of interest.

against it using the likelihood ratio or G^2 test, reported in Table 5. This permits hypothesis tests on the factor groups in model 1 depending only on overall fit rather than on specific parameterizations, such as the Wald tests reported in Table 6.

	df	LOGLIK	DEVIANCE	CHI-SQ	Δdf	Pr(> CHI-SQ)	SIG
Model 1	14	−304.41	608.81				
age	12	−304.71	609.41	0.602	2	0.7402	
proficiency	13	−304.41	608.81	0.004	1	0.9474	
neighborhood	10	−311.55	623.11	14.296	4	0.0064	*
segment	11	−427.14	854.28	245.470	3	< 2.2e-16	***
type	12	−334.63	669.26	60.448	2	7.478e-14	***

TABLE 5. Model comparison using backward elimination from model 1.

Table 5 shows which of the factor groups are considered to affect the phonetic realization of other-language items at a statistically significant level. It reveals that while the fixed effect of mixing type significantly affects the phonetic form that an other-language item will take (as expected by H1), the nature of the segment involved is also statistically influential. In fact, the much greater chi-square values associated with this factor indicate that phonetic realization is affected more by the segment involved than by mixing type (contra H3). The table also demonstrates that the variability is largely independent of extralinguistic considerations: the only extralinguistic factor that significantly affects phonetic realization (the fixed effect of neighborhood) pales in comparison with the linguistic factors of segment and type (contra H2).

Parameters in a logistic regression model are in general not identified without additional assumptions to locate the intercept parameter (Long 1997). We strongly prefer a difference-between-means parameterization for interpretation (Table 6), as this locates the intercept in the center of the data distribution, allowing it to be interpreted as an overall rate of variation in a similar manner to the input value in Varbrul analyses.[28]

RANDOM EFFECTS				
GROUPS	NAME	VARIANCE	SD	
cspkr	(intercept)	0.0687	0.2621	
Cells: 243	cspkr: 25			
FIXED EFFECTS	EST	SE	z-VALUE	Pr(> |z|)	
(intercept)	−0.25595	0.23085	−1.109	0.26754	
type1	−0.17368	0.11862	−1.464	0.14313	
type2	0.84585	0.11299	7.486	7.1e-14	***
segment1	0.90591	0.16550	5.474	4.4e-08	***
segment2	−0.88749	0.15488	5.730	1.0e-08	***
segment3	1.27423	0.13148	9.692	< 2e-16	***
neighborhood1	0.18350	0.22476	0.816	0.41425	
neighborhood2	−0.33319	0.46807	−0.712	0.47657	
neighborhood3	1.34745	0.41350	3.259	0.00112	**
neighborhood4	−0.85002	0.27959	−3.040	0.00236	**
high proficiency	0.01514	0.23886	0.063	0.94947	
age1	−0.09384	0.15942	−0.589	0.55611	
age2	0.09937	0.14574	0.682	0.49535	

TABLE 6. Fixed and random effects of model 1, with Wald tests of parameter values; model AIC: 636.8, BIC: 685.7, G^2: 608.8, df residual: 229.

---

[28] In contrast to Varbrul analyses, however, regardless of parameterization, parameters do not directly express the factors of a factor group. For k factors in a group, there are only $k − 1$ linearly independent parameters in a group.

Profiling the estimates of this model using the scaled zeta values, as recommended by Bates (2010), shows them to be characteristically linear through their respective ranges. Pairwise comparison of the zeta-scores shows only modest linear interdependence between the segment and type factor groups and reveals all other groups to be independent from one another as well as from segment and type,[29] so small changes in one parameter estimate do not unduly distort the model. The model's diagnostics therefore provide us with reasonable confidence in the stability of its structure as estimated and allow us to proceed with their interpretation.

The Wald tests presented in Table 6 are intended to uncover the specific differences within the factor groups identified as significant in Table 5. These reveal that there are significant parameters for each of these factor groups, but that not all of the factors within them are significantly distinct from one another (i.e. neighborhood, where only two of four parameters are significant). To probe further which factors differ significantly, we tested all pairwise differences between factors using the full suite of reference cell parameterizations, Bonferroni-corrected to $p \leq 0.05$ for multiple tests. In Table 7, we report the conditional means for each factor level as computed from the parameters in Table 6, along with the results of the pairwise tests.

NEIGHBORHOOD

	Mont Bleu 1.3474	Basse Ville 0.1835	Vieux Hull −0.3331	West End −0.3477
Basse Ville 0.1835				
Vieux Hull −0.3331				
West End −0.3477	*			
Vanier −0.8500	*	*		

SEGMENT

	P/T/K 1.2742	TH 0.9059	H −0.8875
TH 0.9059			
H −0.8875	*	*	
R −1.2927	*	*	

TYPE

	ATT LWD 0.8458	CS −0.1737
CS −0.1737	*	
NB −0.6721	*	

TABLE 7. Significance tests between categories within each significant factor group; Bonferroni-corrected significant differences are indicated by asterisks in the appropriate cells of each triangular subtable; shading in the left pane identifies the two Quebec neighborhoods.

Table 7 shows that the factor of segment clearly distinguishes the sets {R, H} (disfavoring phonetic integration) and {TH, P/T/K} (favoring it), but does not find the two segments within each set to behave distinctly from the other. The factor group of neighborhood, while significant, only differentiates some of them. The groupings reveal idiosyncrasies that are entirely at odds with the relative intensity of contact each neighborhood instantiates (Table 2 above), contra H2.[30] Even the inner workings of the

---

[29] Most of the factor groups, especially neighborhood, show interdependence with the intercept term. This is an expected consequence of the difference-between-means parameterization, which must locate the intercept differently for any difference in the other parameters estimated.

[30] The two Quebec neighborhoods (Mont Bleu and Vieux Hull, shaded in Table 7) overlap, predictably enough. But when compared to the three Ottawa neighborhoods, any apparent similarity is eroded: Mont Bleu differs statistically from Vanier and West End (although not the intermediate Basse Ville, counterintuitively), while Vieux Hull does not differ from any of them. The Ottawa neighborhoods also diverge from the patterning predicted by their relative level of contact with English. Not only is there no clear distinction between

factor of mixing type defy expectation A/H1, distinguishing ATT LWDs from the other two mixing types, in striking contrast to their morphosyntactic treatment, which segregates CSs.

To better understand why the extralinguistic factors did not emerge as statistically significant, the random effects (which control for the potential influence of individual speakers) are presented in Figure 2, where they are compared against the groupings for age, proficiency, and neighborhood. In these plots, each random effect is plotted together with the upper and lower bounds of its 95% confidence interval (upward- and downward-pointing triangles, respectively). Each speaker belonging to the extralinguistic cohort in the panel (e.g. low English proficiency) is presented left-to-right in order of low-to-high value of the mean of the random effect. Because the scale is the same in each panel, the speakers in each cohort (e.g. low vs. high English-proficiency speakers) can be compared to see if their confidence intervals overlap. For all age and proficiency groups, the variance WITHIN them (as assessed by the vertical range within a panel), including that within each speaker, is greater than that across them (comparing panels), explaining why neither was estimated to significantly influence the variability as a fixed effect. For neighborhood, some groups may appear to be distinct, but the effect is additionally confounded by the small sizes of Vieux Hull and Mont Bleu (each represented by only one speaker). There is therefore no compelling reason to suppose that these factors exert an influence on phonetic realization, confirming our inferences based on the distributions in §3.

In summary, the intraindividual variation overwhelms the interindividual variation in phonetic integration noted in §3, and hence is not well explained by the grouping factors of neighborhood, age, or proficiency. These breakdowns, together with the finding that age and proficiency are not significant as fixed effects, only bolster the result that speakers' extralinguistic characteristics do not clarify the motivations behind phonetic realization.

4.5. MODEL INTERPRETATION. Perhaps most surprising, mixing type, the factor assumed at the outset to be most determinative of propensity to integrate (H1), is revealed by the statistical analysis performed in this section to exert a far more modest effect than expected. Moreover, such effects as do emerge run counter to expectation A, based on robust results in the morphosyntactic domain, that NBs would parallel ATT LWDs (as well as L_R) and that CSs would not be 'integrated' into French at all.

The most important predictor turns out to be segment type.[31] This is quite unexpected under traditional assumptions (our expectation C), because it means that realization will necessarily vary not only across but also within words, as a function of the segments that constitute them. Only four were considered here, but their behaviors contrast significantly (contra H3): two strongly favor integration (P/T/K, TH), while the other two (H, R) are far more likely to retain their L_D form. Regardless of what motivates these differences, the result that they influence the variability so strongly is quite damaging to any claim that phonetic realization is diagnostic of mixing type. Since there are in fact

them and their Quebec counterparts, but two of them also differ statistically from Mont Bleu, while all three overlap with Vieux Hull. Moreover, West End, the neighborhood with the greatest level of contact with English, overlaps with the two Ottawa neighborhoods in which French is more prevalent. Inexplicably, the latter differ statistically from each other.

[31] This robust result is independent of the relative frequency of the segments examined. Additional models run excluding tokens from the low-incidence segments TH and H confirm the significance of this factor as well as its overriding effect.
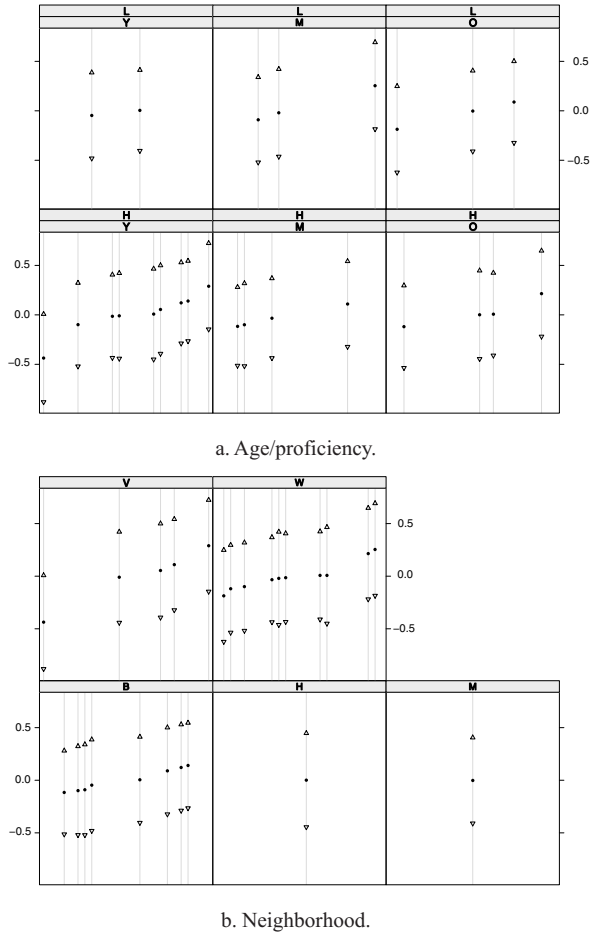
a. Age/proficiency.

b. Neighborhood.

FIGURE 2. Conditional means per individual for model 1 showing relation to grouping factors, with 95%
confidence intervals (upward- and downward-pointing triangles). Age/proficiency: **Y**ounger,
Midrange, **O**lder; **H**igh proficiency, **L**ow proficiency. Neighborhood (in increasing order
of intensity of contact): **M**ont Bleu, Vieux **H**ull, **V**anier, **B**asse Ville, **W**est End.

many more phonetic conflict sites in this language pair than those considered here
(§2.4), we can only speculate on the extent of variability that would have emerged had
more been included. For now, we can confirm that propensity to integrate depends more
on what is being integrated than on any concern to distinguish mixing types.

   Moreover, such regularity as exists among segments is further mediated by the indi-
vidual component of phonetic integration, even among the small cohort who have a
choice in the matter. Results also showed that these individual behaviors are unrelated to
considerations of age or bilingual proficiency. Even the statistical significance of neigh-
borhood of residence (Table 5) is undermined by the result (Table 6) that only two of its
four parameters are statistically distinct. Its influence is further minimized by the wide
range of interindividual variation within neighborhoods (Fig. 2). Table 7 clarifies that
whatever neighborhood effect may be operating owes nothing to intensity of contact at
the local level (see n. 30). Instead we observe shared preferences in neighborhoods with
dissimilar profiles and differentiation between neighborhoods with like characteristics.
We stress that here, as with the other extralinguistic measures, the speakers analyzed rep-
resent only a small proportion of those who actually populate these cohorts, the others

having been excluded from this study for the reasons outlined in §2.2. Observed inter-speaker differences would only intensify were we to consider the entire community, in which, we recall, many speakers have limited or no ability to produce these segments in anything other than L_R. These results underscore the problems associated with diagnosing the language membership of items via their phonetic realization.

5. DISCUSSION AND CONCLUSIONS. Despite pervasive indeterminacy about the precise process underlying loanword adaptation, the default expectation among most scholars is that, barring the odd exception, borrowed words eventually appropriate L_R structure. So strong is this expectation that many simply take for granted that L_D items with L_D phonetics and/or phonology are not borrowed at all, but rather code-switched, while others explicitly invoke such (lack of) integration as a criterion for determining the status of other-language items. Since code-switched and borrowed words (at least established ones) are now known to behave differently in most respects, such assumptions are understandable. But observation of bilingual behavior on the ground suggests that, at the phonetic level, they are often unmet, provoking a great deal of controversy over the role of phonetic and phonological integration in bilingual mixing. Observing that the profusion of literature on this topic nonetheless includes little pertinent information on how borrowed words are treated in actual usage, we sought to contribute to the debate by drawing on a number of lines of empirical evidence to investigate how speakers deploy phonetic integration while (nonce) borrowing, retrieving established loanwords, and code-switching. Major questions include whether they adopt a shared strategy for handling different types of L_D material, as we have seen to be the case (for these same borrowed words in this same community) in the morphosyntactic domain, and, by extension, whether it too qualifies as a diagnostic for identifying and distinguishing these bilingual behaviors.

It is often observed (e.g. Calabrese & Wetzels 2009, Kenstowicz & Suchato 2006) that one of the most intriguing aspects of loanword adaptation is that, at the onset of the borrowing process, there are no precedents guiding which (phonetic or phonological) counterparts to select from the L_R inventory. Borrowing can thus 'reveal aspects of native speakers' knowledge that are not necessarily obvious in data of the native language' (Kang 2011:2258). The role of the initial borrower has therefore assumed particular prominence in the loanword-phonology literature, under the assumption that the strategy s/he selects for integrating the borrowed word will have a crucial (if not a determining) effect on the form the word assumes once it achieves bona fide loanword status. Of course, neither the original borrower nor the initial act of borrowing can be recuperated (even in a laboratory setting, despite valiant efforts). The closest we can come is through the study of nonce borrowing—the locus that offers maximal freedom over whether and how to adapt. Our focus on NBs, contextualized with respect to other mixing types, thus offers an ideal venue to test these assumptions.

As we have seen, however, privileging the adaptation process as it unfolds spontaneously in the community (as opposed to the laboratory or the armchair) is not without its problems. Principal among them is that much of what is conventionally thought of as integration is involuntary. There are at least two reasons for this: (i) many individuals have no choice but to 'integrate' because their inability to produce L_D realizations results, by default, in L_R realizations, and (ii) an unknown portion of the L_D lexical stock regularly adduced as evidence for phonetic adaptation is actually beyond its purview. This is the case of at least some long-attested loanwords, since speakers who no longer recognize them as L_D will inevitably realize them in L_R ('integrate' them) by default, as well as an unknown quantity of (single-word) CSs. These problems are further exacerbated by

the fact that language mixing, despite its salience to community members and linguists, is quite rare in bilingual discourse (e.g. Poplack 2018). We sought to respond to these problems by first locating, in a large sample of bilingual community members, those who engaged in the most code-switching and nonce-borrowing. All were native speakers of French, and by virtue of this fact were of course capable of integrating L_D segments into L_R. To ensure that what we were studying was voluntary integration, we additionally ascertained that they could also produce the L_D segments in L_D. To control for the different integration targets thought to be associated with different types of language mixing, we systematically distinguished ATT LWDs (expected integration rate: full or high, paralleling L_R), multiword CSs (expected integration rate: nil or low, paralleling L_D), and NBs (expected integration rate: paralleling ATT LWDs). We further compared and contrasted both individuals and the extralinguistically defined cohorts they participated in. To home in on the process of integration per se, we retained as diagnostics segments specifically occurring at conflict sites and distinguished among these as well.

What have we learned? Perhaps most surprising is the lack of a systematic strategy for the phonetic treatment, not just of NBs, but also of all types of other-language material. The resulting oscillation between L_D and L_R realizations presented in §3 is a key source of the abiding controversies over the proper avenue for loanword integration (phonetic or phonological), as well as over the identification of language-mixing strategies more generally. To be sure, as observed above, this variability has not gone unrecognized, but the degree and extent uncovered here could not have been predicted based on either the many claims or the few empirical results previously available. Where noted, it has tended to be ascribed to extralinguistic factors like intensity or length of contact, type of transmission, and degree of bilingual proficiency (e.g. Chang 2008, Crawford 2007, Heffernan 2007, Kang 2011, Paradis & LaCharité 2008, Smith 2006, and Yip 2002), among others. This study has demonstrated that, in this community at least, the phonetic treatment of borrowed words shows no such extralinguistic conditioning. Neither intensity of contact at the community level, nor speaker age, nor—even more surprisingly—individual L_D proficiency showed a significant effect. The only extralinguistic factor found by the analyses in §4 to affect variant choice—neighborhood of residence—contributes only a minor effect, and one that is at odds with standard contact-based explanations.

If speakers' disparate integration rates cannot be explained by recourse to such extralinguistic predictors, what DOES condition the variability? Our prediction going in was that mixing type would exert the strongest effect on phonetic realization; the analyses of §4 revealed that its role is not only secondary, however, but also counterintuitive. The strongest determinant of phonetic realization is segment. This unexpected finding means that propensity to integrate a borrowed word depends more on the element to be integrated than on adherence to the grammar operating on that word. Some segments favor L_D realizations; others favor L_R. Here, for instance, a word containing R will be preferentially produced in L_D form whether it is contained within a CS to English or in an old, established loanword like *bar* (dictionary-attested as far back as 1857). P, T, and K, in contrast, will tend to be integrated into French regardless of whether they appear in a borrowed word or in a longer stretch of English. That segment should play any role at all in phonetic integration—let alone the strongest—is most damaging to the assumption that phonetic realization is a necessary or useful adjunct to determining the status of L_D-origin material.

These results are all the more perplexing insofar as they differ from those described in a previous study of the same phenomena in the same community (Poplack et al.

1988). Although those authors also reported phonetic variability at every stage of the borrowing process (including several more stages than we distinguished here), they did detect a modicum of extralinguistic conditioning: phonetic integration appeared to increase as a function of age of attestation of the L_D-origin word and its level of diffusion across the community (1988:73–74, figures 6 and 7). These (widely endorsed) findings have generally been interpreted to mean that phonetic integration increases systematically as the borrowed word achieves bona fide loanword status. It is therefore worth examining why we have been unable to replicate them here.

For one thing, the studies are not entirely comparable in terms of diagnostics, types of borrowed material investigated, and criteria for identifying them. But the crucial distinction resides in an issue raised at the outset: the characteristics of the speakers who provided the data on which the studies are based. Poplack et al. 1988 included the entire Ottawa-Hull sample, nearly five times as many speakers as we have studied here. Most of them—predictably, given the subsampling criteria outlined in §2.2—are far less bilingual than the present participants. Approximately two thirds of the ninety-five non-participants score low on the CEPI (63/95), cite French as their most frequently used language (66/95), and rarely if ever nonce-borrow (58/95) from or code-switch (65/95) into English. The distribution of the subsample analyzed here is basically the inverse, with two thirds (16/25) scoring high on the CEPI, using more English and less French than nonparticipants, and, by definition, engaging in the most CSs and NBs. This, in conjunction with the fact that the L_D material that the ninety-five participants in the 1988 study do employ appears mainly in established loanwords (necessarily, since they rarely if ever nonce-borrow, code-switch, or otherwise speak English) helps explain the Poplack et al. 1988 finding that older and more diffused (i.e. established) loanwords tend to display more phonetic integration than NBs or CSs. It is now apparent that most of the words falling into these categories would have been produced by different cohorts of speakers (see also §2.2), and the ninety-five not studied here were likely 'integrating' their established loanwords involuntarily.

Added to this is the fact that neither words, segments, nor speakers were examined individually in the earlier study (understandably, since it was based on 120 speakers, 19,579 English-origin words, and many, many more segments). This means that the overall integration rates in the 1988 study necessarily conflate disparate elements that we now know do not behave in concert. Once we control systematically for all of these confounding factors, the apparent phonetic trends reported there dissipate. The morphosyntactic results, by contrast, have been replicated many times over. The amalgamated data in the 1988 community study and other work where even less is known about the participants (if any were studied) can thus be interpreted as a cautionary tale about how group trends can obscure individual tendencies (as presaged by Bullock 2009).

In concluding, we offer two important caveats. Even though this study found standard predictors of variability in bilingual contexts not to play a role, we do not believe that we have exhausted the range of possible contributors, linguistic or extralinguistic. Future work should increase efforts to locate the factors conditioning this variability, subject of course to the requirement that they lend themselves to operationalizing and testing.[32] Nor would we claim that the results reported here hold true for every bilingual community. In some contexts, phonetic integration may correlate better with morphosyntactic integration (e.g. Bessett 2016), though it is noteworthy that the very dif-

---

[32] Given the overriding importance of segment, one possibility that comes to mind involves internal structural constraints.

ferent setting and language pair studied by Lev-Ari et al. (2014) displayed the same type of inter- and intraspeaker variability, as well as variability across segments and words. On the contrary, the social circumstances of language mixing should be controlled for in every community study, as should the additional factors we examined here, and still others yet to be uncovered.

We stress, however, that the existence of conditioning, social or linguistic, in no way impeaches the central lesson of this study: speakers do not marshal phonetic integration systematically to distinguish mixing types. Indeed, pending replications of this study in more bilingual communities and language pairs, we submit that our findings are damaging to a number of tropes in the loanword-integration literature. The most important phonetic strategy for handling other-language words—that of treating different segments in different ways—is the greatest detractor from the diagnosticity of phonetic realization. Variability in the phonetic realization of borrowings is not eliminated as established loanwords develop (Chang 2008, Crawford 2007, Kenstowicz & Suchato 2006, LaCharité & Paradis 2005, Paradis & LaCharité 1997, 2008, Peperkamp 2005), but may persist throughout the process. Rates of importation (failure to integrate) do not correlate with level of bilingualism[33] (Chang 2008, Haugen 1950, Paradis & LaCharité 2008, Poplack et al. 1988), nor are meaningful explanations to be found in other extralinguistic escape hatches occasionally appealed to in previous work. This confirms that phonetic and morphosyntactic integration can proceed independently; the hypothesis that ATT LWDs and NBs are systematically related is thus supported at the morphosyntactic level, but emphatically not at the phonetic one. These and the other findings presented here all prove that speakers do not avail themselves of phonetics strategically to distinguish other-language phenomena. Consequently, linguists using this as an unconsidered criterion for identifying language-mixing strategies do so at their peril. Phonetic integration of borrowed material should not simply be assumed; it must be established segment by segment, word by word, speaker by speaker, and community by community.

<div align="center">APPENDIX</div>

A1. DATA. We present in Table A1 a complete cross-tabulation of the data on which this study is based. Values are indicated as number of observations of French:English phonetic form. Dashes indicate no observations of either for a speaker in a given context (i.e. 0:0).

As indicated by Table A1, the data are quite sparse. Of the 300 total cells in Table A1, fifty-seven are unobserved for both French and English (0:0, represented with —) (i.e. environments that did not exist in a given speaker's data, such as a CS containing the segment H), while eighty-nine more are unobserved for French (0 before the colon) and fifty-four for English (0 after the colon) (i.e. contexts where the speaker made use of only one variant). Similarly, only eight of the 300 cells have greater than ten observations, eighty-two have between five and ten (inclusive), and 153 have fewer than five, not including the unobserved cells. As this exhausts the available bilingual corpus (see the discussion of Table 1 in §2.4), some potential research questions cannot be answered. For example, one referee asked to see per-speaker random slopes for type within each of the four segments. Not only does this saturate the model, but some segments, such as TH, clearly lack the necessary observations: 80% of the speakers do not have this segment in their NBs, and 52% ($N = 13$) lack it in their ATT LWDs. At the very least, using classical methods of estimation as in lme4, one can anticipate that complex models will fail to converge because of sparsity and/or data imbalance.

A2. RANDOM SLOPES. Because of the sparsity of our data, we strongly prefer a simpler model with per-speaker random intercepts and no random slopes. However, one might for a priori reasons prefer the model to have, for instance, per-speaker random slopes for type, reasoning that any cross-speaker effect of type needs to account for whether speakers taken individually exhibit an effect. We find this to push the limits of our data. Models without random slopes are also more generous to the hypotheses we wish to argue against—for example, that age and/or proficiency bear on phonetic integration. Including random slopes

---

[33] Except, of course, when monolinguals are factored in.

Neigh	Age	Prof	Spkr	CODE-SWITCH				ATTESTED LOANWORD				NONCE BORROWING			
				TH	H	P/T/K	R	TH	H	P/T/K	R	TH	H	P/T/K	R
B	Y	H	f	2:2	3:2	3:1	0:3	3:0	0:1	10:2	5:5	—	—	4:1	1:4
			g	1:1	1:4	4:1	0:5	—	0:8	5:0	0:1	—	0:1	4:0	1:2
			j	5:0	1:3	4:0	0:4	—	0:1	5:0	1:10	—	—	5:0	2:5
		L	h	4:1	3:2	5:0	0:4	2:0	2:1	4:3	1:9	—	0:1	1:1	1:2
			i	1:3	1:1	—	0:5	—	1:0	2:1	5:5	1:0	—	1:2	0:2
	M	H	k	2:2	1:4	0:1	0:5	5:0	3:0	4:0	3:2	2:2	0:2	2:2	0:19
		L	l	5:0	—	4:3	0:5	3:0	1:0	5:0	0:6	—	—	1:1	—
	O	L	m	3:2	2:0	3:1	0:4	1:0	0:1	4:2	5:8	—	1:0	2:0	2:4
H	O	H	x	0:3	1:3	1:4	1:3	—	—	2:0	5:0	—	—	0:2	0:2
M	O	L	y	1:0	2:0	6:0	0:4	—	4:2	22:0	6:5	0:1	—	3:0	3:4
V	Y	H	a	1:1	1:1	0:1	0:2	2:0	3:5	6:1	3:3	—	—	1:0	0:3
			c	0:3	—	1:2	0:2	0:3	0:5	0:4	0:4	—	0:1	1:2	0:4
			d	2:2	0:3	4:0	0:4	—	0:6	0:3	2:5	1:0	—	1:2	0:4
			e	0:5	0:4	4:0	0:5	3:0	0:1	1:1	2:0	—	—	0:1	0:3
	M	H	b	3:2	2:3	2:2	0:5	—	—	9:2	1:2	—	—	0:1	0:4
W	Y	H	n	2:3	—	4:0	0:5	—	0:8	5:2	4:1	—	0:1	1:1	0:4
			s	4:0	0:5	0:1	0:5	—	—	1:0	2:0	—	—	0:3	0:3
	M	H	o	1:3	0:4	4:0	0:4	0:1	—	5:0	3:3	1:0	0:3	4:5	1:19
			v	0:3	1:2	3:2	0:5	—	0:5	4:0	3:4	—	—	3:2	0:4
		L	t	1:1	0:3	1:0	0:2	5:0	0:1	5:1	1:5	—	0:1	1:0	0:4
			u	—	2:0	3:1	0:4	2:0	5:2	6:0	1:1	—	—	0:1	0:4
	O	H	q	1:2	1:2	3:1	0:4	—	3:6	3:2	1:3	—	0:1	3:1	1:5
			r	2:1	0:2	3:0	2:3	—	0:1	6:0	3:4	—	—	2:2	0:2
			w	4:1	0:5	2:0	0:4	2:4	2:2	2:3	3:3	—	—	1:2	0:5
		L	p	3:1	0:4	0:3	0:5	1:0	1:3	2:2	2:1	—	0:1	0:2	0:5

TABLE A1. Complete cross-tabulation of data by speaker (within neighborhood, age, and proficiency), mixing type (CS, ATT LWD, and NB), and segment (TH, H, P/T/K, and R).

raises the risk of type II error (Baayen et al. 2015, Matuschek et al. 2017), so a stronger argument that no effect is found is made by a simpler model in which these factors have a better opportunity to be associated with some of the observed variation.

Models with random slopes may be fitted to our data by adopting a Bayesian approach. The R package brms (Bürkner 2017) fits a variety of random-slope models to our data. One such model, presented in Table A2, adds to model 1 (given in 10 above) a term for per-speaker effects of type (0 + type ∥ spkr) and one for segment (1 + seg ∥ spkr), the double-bar notation preventing the estimation of correlation parameters among the random effects. This model, model 3 (the reason for this numbering to be made clear below), addresses a potential criticism of our interpretation of the effect of mixing type and segment in model 1, as the effect of type in model 3 must take into account interindividual differences in the effect of type.

Since model 3 employs Bayesian assumptions, its parameter report differs from that of Table 6 above; estimates are presented with their standard error, lower and upper bounds of a 95% confidence interval of the estimate, effective sample size, and R-hat. The last two columns reflect the parameter-wise sampling performance of the sampler; effective sample sizes in the hundreds to thousands are considered acceptable for inference, and R-hat values close to 1 indicate acceptable convergence. Inspection of the confidence intervals for inclusion of zero is roughly equivalent to a two-tailed Wald hypothesis test. We note that these parameters, indicated by asterisks (*) in Table A2, match those identified as significant in model 1 in Table 6—except for neigh4, which is marginal (indicated by a dot ·) in that the upper tail of its parameter distribution just crosses zero. Hence, none of our interpretations of model 1 need to change for model 3, particularly since the interpretation of neighborhood was negative as concerned the hypothesis regarding relative exposure to contact.

It must be stressed that convergence under Monte Carlo methods is very different from that under classical methods as in lme4, and successful convergence does not entail that the model is not overspecified. In fact,

GROUP-LEVEL EFFECTS: SPEAKER ($N = 25$)

	EST	SE	CI LOWER	CI UPPER	EFF SAMPLE	R-HAT	
sd(intercept)	0.36	0.23	0.02	0.88	640	1.00	
sd(seg1)	0.54	0.30	0.03	1.16	770	1.00	
sd(seg2)	0.61	0.30	0.06	1.21	783	1.00	
sd(seg3)	0.42	0.24	0.02	0.90	685	1.00	
sd(typeC)	0.29	0.21	0.01	0.79	1,109	1.00	
sd(typeL)	1.31	0.37	0.71	2.12	1,170	1.00	
sd(typeN)	0.64	0.41	0.04	1.59	1,040	1.00	

POPULATION-LEVEL EFFECTS

	EST	SE	CI LOWER	CI UPPER	EFF SAMPLE	R-HAT	
(intercept)	−0.43	0.37	−1.17	0.32	2,145	1.00	
neigh1	0.50	0.36	−0.21	1.23	2,050	1.00	
neigh2	−1.12	0.71	−2.58	0.23	1,888	1.00	
neigh3	1.80	0.72	0.42	3.19	1,709	1.00	*
neigh4	−0.81	0.43	−1.67	0.05	1,954	1.00	.
age1	−0.07	0.25	−0.57	0.43	2,178	1.00	
age2	−0.10	0.25	−0.60	0.41	1,881	1.00	
high proficiency	0.10	0.38	−0.63	0.86	1,933	1.00	
type1	−0.25	0.18	−0.61	0.11	1,930	1.00	
type2	1.05	0.24	0.58	1.54	1,796	1.00	*
seg1	1.03	0.22	0.58	1.48	2,524	1.00	*
seg2	−0.98	0.23	−1.44	−0.55	2,927	1.00	*
seg3	1.41	0.18	1.06	1.77	2,794	1.00	*

TABLE A2. Mixed-effects model 3 with per-speaker random slopes for type.

	WAIC				LOOIC			
MODEL	EST	SE	CI LOWER	CI UPPER	EST	SE	CI LOWER	CI UPPER
3	604.3	26.6	552.2	656.4	627.3	28.5	571.4	683.2
2	603.3	29.1	569.8	660.3	611.3	29.9	552.7	669.9
1	647.5	33.9	580.9	713.7	650.0	34.2	583.0	717.0

TABLE A3. Model fit information criteria for model 3, model 2, and model 1 as estimated using brms.

other problematic indicators emerge for random-slope models of our data. Specifically, model comparison (e.g. of model 1 and model 3) requires assessing model fit, but the available measures of fit, such as the WIDELY APPLICABLE INFORMATION CRITERION (WAIC) and the LEAVE-ONE-OUT INFORMATION CRITERION (LOOIC), flag multiple cells that are problematic for the relevant assumptions (as many as sixty-three cells or 26% of the data with WAIC and twenty-three cells or 9% for LOOIC). Furthermore, when computed, confidence intervals of the information criteria for model 1 and random-slope models such as model 3 overlap substantially, leaving no obvious justification to prefer the clearly more complex random-slope models. This is demonstrated in Table A3, where model 1 and model 3 are compared alongside model 2, a model with a random-slope component for type but not for segment. For each of these models, the confidence intervals of whichever fit criterion is chosen overlap so much that there is no clear fit-based reason to prefer either of the random-slope models—model 2 and model 3—over the simple random-intercept model 1. The fit criteria of these models are also suspect under the assumptions for estimating WAIC or LOOIC, whereas for model 1 no such problem arises. For these reasons, we do not consider random-slope models of our data to be suitable for interpretation, and consequently we adopt model 1. While not everyone may agree with this analytical judgment, the foregoing should demonstrate that it does not materially change the central interpretations of this study.

REFERENCES

ADALAR, NEVIN, and SALI TAGLIAMONTE. 1998. Borrowed nouns; bilingual people: The case of the 'Londrali' in Northern Cyprus. *International Journal of Bilingualism* (Special issue: *Instant loans, easy conditions: The productivity of bilingual borrowing*, ed. by Shana Poplack and Marjory Meechan) 2.139–59. DOI: 10.1177/136700699800200 202.

AGRESTI, ALAN. 1996. *An introduction to categorical data analysis*. New York: Wiley.

Baayen, R. Harald; Douglas J. Davidson; and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59.390–412. DOI: 10.1016/j.jml.2007.12.005.

Baayen, R. Harald; Shravan Vasishth; Douglas Bates; and Reinhold Kliegl. 2015. Out of the cage of shadows. arXiv:1511.03120 [stat.AP]. Online: https://arxiv.org/abs /1511.03120v1.

Barber, Katherine (ed.) 2001. *The Canadian Oxford dictionary*. 2nd edn. Toronto: Oxford University Press.

Barr, Dale J.; Roger Levy; Christoph Scheepers; and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68.255–78. DOI: 10.1016/j.jml.2012.11.001.

Bates, Douglas. 2010. lme4: Mixed-effects modeling with R. Madison: University of Wisconsin, ms. Online: http://lme4.r-forge.r-project.org/book/front.pdf.

Bates, Douglas; Martin Maechler; Ben Bolker; and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1–48. DOI: 10.18637/jss.v067.i01.

Bessett, Ryan M. 2016. The role of phonology in discerning lone other-language items as borrowing or codeswitches. Paper presented at New Waves of Analyzing Variation (NWAV) 45, Vancouver, BC.

Boersma, Paul, and Silke Hamann. 2009. Loanword adaptation as first-language phonological perception. *Loanword phonology*, ed. by Andrea Calabrese and W. Leo Wetzels, 11–58. Amsterdam: John Benjamins.

Boersma, Paul, and David Weenink. 2016. Praat: Doing phonetics by computer. Version 6.0.15. Online: http://www.praat.org/.

Broselow, Ellen. 2009. Stress adaptation in loanword phonology: Perception and learnability. *Phonology in perception*, ed. by Paul Boersma and Silke Hamann, 191–234. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219234.191.

Budzhak-Jones, Svitlana. 1998. *Single-word incorporations in Ukrainian-English bilingual discourse: Little things mean a lot*. Ottawa: University of Ottawa dissertation.

Budzhak-Jones, Svitlana, and Shana Poplack. 1997. Two generations, two strategies: The fate of bare English-origin nouns in Ukrainian. *Journal of Sociolinguistics* 1.225–58. DOI: 10.1111/1467-9481.00013.

Bullock, Barbara E. 2009. Phonetic reflexes of code-switching. *The Cambridge handbook of linguistic code-switching*, ed. by Barbara E. Bullock and Almeida Jacqueline Toribio, 163–81. Cambridge: Cambridge University Press. DOI: 10.1017/CBO978051 1576331.011.

Bullock, Barbara E., and Almeida Jacqueline Toribio. 2009. Trying to hit a moving target: On the sociophonetics of code-switching. *Multidisciplinary approaches to code-switching*, ed. by Ludmila Isurin, Donald Winford, and Kees de Bot, 189–206. Amsterdam: John Benjamins. DOI: 10.1075/sibil.41.12bul.

Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80.1–28. DOI: 10.18637/jss.v080.i01.

Calabrese, Andrea. 2009. Perception, production and acoustic inputs in loanword phonology. In Calabrese & Wetzels, 59–115.

Calabrese, Andrea, and Leo Wetzels. (eds.) 2009. *Loan phonology*. Amsterdam: John Benjamins.

Caramazza, Alfonso, and Grace H. Yeni-Komshian. 1974. Voice onset time in two French dialects. *Journal of Phonetics* 2.239–45. DOI: 10.1016/S0095-4470(19)312744.

Centre national de la recherche scientifique. 2004. *Trésor de la langue française : Dictionnaire de référence des XIXe et XXe siècles en 16 volumes*. Lorraine: CNRS & Université de Lorraine. Online: http://atilf.atilf.fr.

Chang, Charles B. 2008. Phonetics vs. phonology in loanword adaptation: Revisiting the role of the bilingual. *Berkeley Linguistics Society* 34.43–54. DOI: 10.3765/bls.v34i1 .3557.

Crawford, Clifford. 2007. The role of loanword diffusion in changing adaptation patterns: A study of coronal stops in Japanese borrowings. *Working Papers of the Cornell Phonetics Laboratory* 16.32–56. Online: http://conf.ling.cornell.edu/plab/paper/Crawford _preprint.pdf.

DE VILLIERS, MARIE-ÉVA. 1997. *Le Multi des jeunes : Dictionnaire de la langue française*. Montreal: Québec/Amérique.

DION, NATHALIE. 2003. L'effacement du que en français canadien : Une étude en temps réel. Ottawa: University of Ottawa M.A. thesis.

DIONNE, NARCISSE-EUTROPE. 1909. *Le parler populaire des canadiens français*. Quebec: J. P. Garneau.

ELSIG, MARTIN. 2009. *Grammatical variation across space and time: The French interrogative system*. Amsterdam: John Benjamins.

EZE, EJIKE. 1997. *Aspects of language contact: A variationist perspective on code switching and borrowing in Igbo-English bilingual discourse*. Ottawa: University of Ottawa dissertation. DOI: 10.20381/ruor-8187.

GELMAN, ANDREW, and JENNIFER HILL. 2007. *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

GHAFAR SAMAR, REZA, and MARJORY MEECHAN. 1998. The null theory of code-switching versus the nonce borrowing hypothesis: Testing the fit in Persian-English bilingual discourse. *International Journal of Bilingualism* 2.203–20. DOI: 10.1177/136700699800200205.

HAUGEN, EINAR. 1950. The analysis of linguistic borrowing. *Language* 26.210–31. DOI: 10.2307/410058.

HEFFERNAN, KEVIN. 2007. The role of phonemic contrast in the formation of Sino-Japanese. *Journal of East Asian Linguistics* 16.61–86. DOI: 10.1007/s10831-006-9007-8.

JACOBS, HEIKE, and CARLOS GUSSENHOVEN. 2000. Loan phonology: Perception, salience, the lexicon and optimality theory. *Optimality theory: Phonology, syntax, and acquisition*, ed. by Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 193–209. Oxford: Oxford University Press.

KABANO, ALPHONSE. 2004. *Étude de la grammaticalisation en contexte restrictif et variable : Le pronom personnel sujet en français parlé*. Ottawa: University of Ottawa dissertation.

KANG, YOONJUNG. 2010. The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology* 27.225–53. DOI: 10.1017/S09526757110000114.

KANG, YOONJUNG. 2011. Loanword phonology. *The Blackwell companion to phonology, vol. 4: Phonological interfaces*, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 2258–82. London: Wiley-Blackwell.

KENSTOWICZ, MICHAEL. 2005. The phonetics and phonology of loanword adaptation. *Proceedings of the first European Conference on Korean Linguistics*, ed. by Sang-Jik Rhee, 17–32. Seoul: Hankook.

KENSTOWICZ, MICHAEL, and ATIWONG SUCHATO. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116.921–49. DOI: 10.1016/j.lingua.2005.05.006.

LACHARITÉ, DARLENE, and CAROLE PARADIS. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36.223–58. DOI: 10.1162/0024389053710666.

LEALESS, ALLISON V. 2014. *« J'ai tout le temps eu de misère » : A variationist study of adverb placement in Quebec French*. Ottawa: University of Ottawa dissertation. DOI: 10.20381/ruor-3747.

LEROUX, MARTINE, and LIDIA-GABRIELA JARMASZ. 2006. A study about nothing: Null subjects as a diagnostic of convergence between English and French. *University of Pennsylvania Working Papers in Linguistics (Papers from NWAV 34)* 12:2. Online: https://repository.upenn.edu/pwpl/vol12/iss2/2.

LEV-ARI, SHIRI, and SHARON PEPERKAMP. 2014. An experimental study in the role of social factors in language change: The case of loanword adaptations. *Laboratory Phonology* 5.379–401. DOI: 10.1515/lp-2014-0013.

LEV-ARI, SHIRI; MARCELA SAN GIACOMO; and SHARON PEPERKAMP. 2014. The effect of domain prestige and interlocutors' bilingualism on loanword adaptations. *Journal of Sociolinguistics* 18.658–84. DOI: 10.1111/josl.12102.

LONG, J. SCOTT. 1997. *Regression models for limited and categorical dependent variables*. Thousand Oaks, CA: Sage.

MATUSCHEK, HANNES; REINHOLD KLIEGL; SHRAVAN VASISHTH; HARALD BAAYEN; and DOUGLAS BATES. 2017. Balancing type I error and power in linear mixed models. *Journal of Memory and Language* 94.305–15. DOI: 10.1016/j.jml.2017.01.001.

MEECHAN, MARJORY, and SHANA POPLACK. 1995. Orphan categories in bilingual discourse: A comparative study of adjectivization strategies in Wolof-French and Fongbe-French. *Language Variation and Change* 7.169–94. DOI: 10.1017/S0954394500000971.

MENEY, LIONEL. 1999. *Dictionnaire québécois français*. 1st edn. Montréal: Guérin.

MUSTAFAWI, EIMAN. 2002. Lone English-origin nouns in Arabic: Codeswitches or borrowings? *Proceedings of the 2002 annual conference of the Canadian Linguistic Association*, 219–31. Online: http://homes.chass.utoronto.ca/~cla-acl/2002/Mustafawi_2002 .pdf.

NAULT, KARIN. 2002. 'Ja, solche Sprache-Ø, jetzt geht's schon los, gell?': Zeroing in on German plurals: A variationist approach to German/English language contact. Ottawa: University of Ottawa M.A. thesis.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33.379–430. DOI: 10.1017/S0022226797006 786.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 2008. Apparent phonetic approximation: English loanwords in Old Quebec French. *Journal of Linguistics* 44.87–128. DOI: 10.1017 /S0022226707004963.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 2012. The influence of attitude on the treatment of interdentals in loanwords: Ill-performed importations. *Catalan Journal of Linguistics* 11.97–126. DOI: 10.5565/rev/catjl.12.

PEPERKAMP, SHARON. 2005. A psycholinguistic theory of loanword adaptations. *Berkeley Linguistics Society* 30.341–52. DOI: 10.3765/bls.v30i1.919.

PEPERKAMP, SHARON, and EMMANUEL DUPOUX. 2003. Reinterpreting loanword adaptations: The role of perception. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, 367–70. Online: https://www.internationalphonetic association.org/icphs-proceedings/ICPhS2003/p15_0367.html.

PEPERKAMP, SHARON; INGA VENDELIN; and KIMIHIRO NAKAMURA. 2008. On the perceptual origins of loanword adaptations: Experimental evidence from Japanese. *Phonology* 25.129–64. DOI: 10.1017/S0952675708001425.

PINHEIRO, JOSÉ C., and DOUGLAS M. BATES. 2000. *Mixed-effects models in S and S-Plus*. Berlin: Springer.

POIRIER, CLAUDE. 1998. *Dictionnaire historique du français québécois*. Laval: Trésor de la langue française au Québec.

POPLACK, SHANA. 1985. Contrasting patterns of code-switching in two communities. *Methods V: Papers from the V International Conference on Methods in Dialectology*, ed. by Henry J. Warkentyne, 363–85. Victoria, BC: University of Victoria.

POPLACK, SHANA. 1988a. Conséquences linguistiques du contact de langues : Un modèle d'analyse variationniste. *Langage et société* 43.23–48. DOI: 10.3406/lsoc.1988.3000.

POPLACK, SHANA. 1988b. Language status and language accommodation along a linguistic border. *Language spread and language policy: Issues, implications, and case studies* (GURT 87), ed. by Peter H. Lowenberg, 90–118. Washington, DC: Georgetown University Press.

POPLACK, SHANA. 1989. The care and handling of a mega-corpus. *Language change and variation*, ed. by Ralph W. Fasold and Deborah Schiffrin, 411–51. Amsterdam: John Benjamins.

POPLACK, SHANA. 2015. Norme prescriptive, norme communautaire et variation diaphasique. *Variations diasystématiques et leurs interdépendances* (Travaux de linguistique romane: Série TraLiRo), ed. by Kirsten Jeppesen Kragh and Jan J. Lindschouw, 293–319. Zürich: Société de Linguistique Romane.

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

POPLACK, SHANA, and NATHALIE DION. 2012. Myths and facts about loanword development. *Language Variation and Change* 24.279–315. DOI: 10.1017/S09543945120001 8X.

POPLACK, SHANA, and MARJORY MEECHAN. 1995. Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, ed. by Lesley Milroy and Pieter Muysken, 199–232. Cambridge: Cambridge University Press. DOI: 10.1017 /CBO9780511620867.010.

POPLACK, SHANA, and MARJORY MEECHAN (eds.) 1998a. *Instant loans, easy conditions: The productivity of bilingual borrowing*. (Special issue: *International Journal of Bilingualism* 2(2).)

POPLACK, SHANA, and MARJORY MEECHAN. 1998b. How languages fit together in codemixing. *International Journal of Bilingualism* 2.127–38. DOI: 10.1177/136700699800200201.

POPLACK, SHANA; DAVID SANKOFF; and CHRISTOPHER MILLER. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26.47–104. DOI: 10.1515/ling.1988.26.1.47.

POPLACK, SHANA; LOTFI SAYAHI; NAHED MOURAD; and NATHALIE DION. 2015. An exception to the rule? Lone French nouns in Tunisian Arabic. *University of Pennsylvania Working Papers in Linguistics (Selected papers from NWAV 43)* 21(2):20. Online: https://repository.upenn.edu/pwpl/vol21/iss2/20.

POPLACK, SHANA, and ANNE ST-AMAND. 2007. A real-time window on 19th-century vernacular French: The *Récits du français québécois d'autrefois*. *Language in Society* 36. 707–34. DOI: 10.1017/S0047404507070662.

POPLACK, SHANA, and DANIELLE TURPIN. 2009. Does the *Futur* have a future in (Canadian) French? *Probus* 11.133–64. DOI: 10.1515/prbs.1999.11.1.133.

POPLACK, SHANA; LAUREN ZENTZ; and NATHALIE DION. 2012. Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition* 15.203–25. DOI: 10.1017/S1366728911000204.

ROBERT, PAUL. 1977. *Petit Robert : Dictionnaire alphabétique et analogique de la langue française*. 1st edn. Paris: Société du nouveau Littré.

ROBERT, PAUL. 1986. *Petit Robert : Dictionnaire alphabétique et analogique de la langue française*. 2nd edn. Paris: Société du nouveau Littré.

ROBERT, PAUL. 2004. *Petit Robert: Dictionnaire alphabétique et analogique de la langue française*. Revised edn. Paris: Dictionnaires Le Robert.

ROSE, YVAN, and KATHERINE DEMUTH. 2006. Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua* 116.1112–39. DOI: 10.1016/j.lingua.2005.06.011.

SANKOFF, DAVID; SHANA POPLACK; and SWATHI VANNIARAJAN. 1990. The case of the nonce loan in Tamil. *Language Variation and Change* 2.71–101. DOI: 10.1017/S0954394500000272.

SILVERMAN, DANIEL. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9.289–328. DOI: 10.1017/S0952675700001627.

SMITH, JENNIFER. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. *Japanese/Korean linguistics 14*, ed. by Timothy J. Vance and Kimberly A. Jones, 63–74. Stanford, CA: CSLI Publications.

STAN DEVELOPMENT TEAM. 2018. RStan: The R interface to Stan. R package version 2.18.2. Online: http://mc-stan.org.

TORRES CACOULLOS, RENA, and JESSI ELANA AARON. 2003. Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change* 15.287–326. DOI: 10.1017/S0954394503153021.

TURPIN, DANIELLE. 1995. «Le français, c'est le *last frontier* » : La structure nominale dans le discours bilingue français-anglais. Ottawa: University of Ottawa M.A. thesis. DOI: 10.20381/ruor-8285.

VENDELIN, INGA, and SHARON PEPERKAMP. 2006. The influence of orthography on loanword adaptations. *Lingua* 116.996–1007. DOI: 10.1016/j.lingua.2005.07.005.

WHEELER, SUSAN M. 1987. Code switching and borrowing in a Finnish-English bilingual situation. Ottawa: University of Ottawa M.A. thesis. DOI: 10.20381/ruor-10612.

WILLIS, LAUREN A. 2000. *Être ou ne plus être*: Auxiliary alternation in Ottawa-Hull French. Ottawa: University of Ottawa M.A. thesis. DOI: 10.20381/ruor-16273.

YIP, MOIRA. 2002. Perceptual influences in Cantonese loanword phonology. *Journal of the Phonetic Society of Japan* 6.4–21. DOI: 10.24467/onseikenkyu.6.1_4.

YIP, MOIRA. 2006. The symbiosis between perception and grammar in loanword phonology. *Lingua* 116.950–75. DOI: 10.1016/j.lingua.2005.05.007.

Poplack
Department of Linguistics
University of Ottawa
401-70 Laurier Avenue East
Ottawa, ON K1N 6N5, Canada
[spoplack@uottawa.ca] (Poplack)
[suzanne.robillard@uottawa.ca] (Robillard)
[nathalie.dion@uottawa.ca] (Dion)
[paolillo@indiana.edu] (Paolillo)