Contact-induced grammatical change: A cautionary tale

Chapter · January 2010

CITATIONS
78

READS
449

2 authors, including:

Shana Poplack
University of Ottawa
122 PUBLICATIONS **7,570** CITATIONS

SEE PROFILE

Language and Space

An International Handbook of Linguistic Variation

Volume 1: Theories and Methods

Edited by

Peter Auer and Jürgen Erich Schmidt

Offprint

Poplack, Shana & Levey, Stephen. 2010. Contact-induced grammatical change. In Auer, Peter & Schmidt, Jürgen Erich (eds.), Language and Space – An international handbook of linguistic variation: Volume 1 – Theories and methods. Berlin: Mouton de Gruyter. 391-419.

De Gruyter Mouton

IV. Structure and dynamics across language spaces

22. Contact-induced grammatical change: A cautionary tale

1. Introduction

When the space between two languages – or their speakers – is reduced enough to result in *contact,* a variety of outcomes may ensue. Most prominent by far is *change.* So ubiquitous has the association between contact and change become that it is invoked to account for a wide array of developments that differ from some (typically ill-defined) norm. This is the case even where the candidate development already exists in the recipient language, where no other-language form is visible, and where the transition period is as short as a single generation. And despite the general consensus that grammatical structure is particularly resistant to transfer, reports of dramatic structural changes in languages known to have been in contact are legion.

In this article we suggest that much of the evidence brought to bear on contact-induced change – diachronic as well as synchronic – either fails to demonstrate that change has occurred, and/or if it has, that it is the product of contact and not internal evolution. These issues, together with the possibility that the inherent variability characteristic of all spoken language may have been mistaken for change, need to be resolved before a contact explanation can be justified.

In the ensuing sections we first provide a brief overview of some of the more salient types of change reported to result from contact. We then review the kinds of confounding evidence that plagues the study of change, contact-induced or otherwise, and propose criteria for establishing its existence empirically. In section 3 we detail a method, combining the machinery of variationist sociolinguistics with the comparative method of historical linguistics, capable of identifying change and ascertaining its source. We then present a series of case studies illustrating the utility of this approach and conclude with some observations about the implications of our findings for understanding the existence, nature and extent of contact-induced change. Our focus is on the grammar, but the considerations we present also apply to other types of change, with the exception of lexical borrowing.

2. Background

Contact-induced structural change has long been a subject of debate among linguists, due in no small part to the primacy traditionally accorded the internal evolution of language (e.g., Lass 1980; Martinet 1960). This was traditionally assumed to be regular and rule-governed, while externally-motivated changes could be arbitrary, unpredictable and idiosyncratic. External explanations were thus often invoked as a last resort, when a more satisfactory account was not available (Romaine 1988: 349).

Early opinions ranged from outright denial of the possibility of a mixed language (e.g., Müller 1871: 86; Whitney 1881) to the claim that there are no completely *unmixed* languages (Schuchardt 1884: 5). But most of those who admitted the possibility of contact-induced change did so cautiously and under specific conditions. Haugen's comment that "[on] those relatively rare occasions when bound morphemes are actually transferred from one language to another, they are such as fit readily into a preexistent category in the recipient language" (1954: 386), is but one example. Similarly, Jakobson (1962: 241) argued that "a language accepts foreign structural elements only when they correspond to its own developmental tendencies". Notwithstanding such caveats, major structural changes said to result from language contact continue to be reported.

Among the more spectacular changes of the past figure the borrowing of Turkish inflectional suffixes into dialects of Asia Minor Greek (Dawkins 1916), the incorporation of the entire Spanish preposition system into Malinche Mexicano (Field 2005: 352), the borrowing of Russian finite verb morphology into Mednyj Aleut (Thomason 2001), the development of Michif from French noun phrases and Cree verb phrases (e.g., Bakker 1997; Bakker and Papen 1997) and the (in)famous convergence of Kupwar Urdu, Marathi and Kannada into "a single surface syntactic structure" (Gumperz and Wilson 1971: 155).

After reviewing many such changes, as well as a variety of theories of what constrains them, Thomason was led to reject the idea that categorical constraints against contact-induced change could be enunciated. Instead she proposed that under the right social and linguistic conditions,

> [...] anything goes, including structural borrowing that results in major typological changes in the borrowing language. In phonology, loss or attrition of entire phonetic and/or phonological categories in native words and of all kinds of morphophonemic rules. In syntax, sweeping changes in such features as word order, relative clauses, negation, coordination, subordination, comparison, and quantification. In morphology, typologically disruptive changes such as the replacement of flexional by agglutinative morphology or vice-versa, the addition or loss of morphological categories that do not match in source and borrowing languages, and the wholesale loss or addition of agreement patterns. (Thomason 2001: 71)

At the time of this writing, Thomason's (and Schuchardt's) views enjoy wide support. The implicit assumption is that change is an almost inevitable result of language contact (e.g., Appel and Muysken 1987: 154; Bynon 1977: 240; Harris and Campbell 1995: 149, among many others; although see Thomason 2007 for discussion of deliberate resistance to contact-induced change). To be sure, some scholars continue to caution against promiscuous inferences, particularly in the absence of robust sociohistorical and linguistic evidence (e.g., Dorian 1993; Mougeon, Nadasdi and Rehner 2005). And Heine and Kuteva (2005: 8) observe that contact is often invoked without satisfactorily ruling out

other causal mechanisms. Nonetheless, the onus seems to be on the analyst to justify why change did *not* occur in an apparently favorable situation:

> But since [the above] predictions are robust – that is, they are valid in the great majority of cases that have been described in the literature – any *violation* should provide interesting insights into social, and, to a lesser extent, linguistic determinants of contact-induced change. (Thomason 2001: 71, emphasis ours)

And in fact, the kinds of changes enumerated above are reported in contemporary contact situations as well. In the least conspicuous cases, an existing minor-use pattern is claimed to be developing into a major-use pattern via overextension of recipient-language options into new contexts (e.g., Heine and Kuteva 2005: 44–62; Klein 1980; Silva-Corvalán 1994a: 167; Toribio 2004: 167). Thus Toribio (2004: 170) states that when the "bilingual mode" of two Spanish-English bilinguals was experimentally induced, their use of overt subject pronouns, while not ungrammatical in Spanish, reflected the influence of English, a non-pro-drop language. Savić (1995: 487, cited in Heine and Kuteva 2005: 69) reaches similar conclusions in relation to slightly elevated uses of optional overt subject and object NPs in the discourse of Serbian-English bilinguals. Heine and Kuteva also discuss an increase of overt pronouns among Russian and Hungarian bilinguals, again on the model of English (2005: 68), as do Otheguy, Zentella and Livert (2007) for New York City Spanish and Backus (2005: 333) for Turkish. The same process has been said to initiate word order changes, as in the shift from the Spanish majority pattern of postposed adjectival placement (*la persona más importante*) to the pre-nominal adjectival placement (*la más importante persona*) characteristic of English (Toribio 2004: 167).

In another type of change, recipient-language morphemes may encode a completely new grammatical category, as in Tariana-Portuguese bilinguals' use of interrogative pronouns (e.g., *kwana* 'who') to mark relative and complement clauses on the model of Portuguese, or of Portuguese lexemes to express the Tariana evidentiality marking system (Aikhenvald 2002: 182–183, 315–316).

Other contemporary cases of convergence involve outright violation of recipient-language grammar. Clyne (2003: 132) discusses overgeneralization by Dutch-English and, to a lesser extent, German-English bilinguals of SVO word order in contexts where SOV and VSO are prescribed in the non-contact varieties.

Apart from extension of an existing option, addition of a totally new category and violation of recipient-language rules, contact is widely believed to accelerate attrition and loss of grammatical features in one language with no equivalent in the other (e.g., Seliger and Vago 1991). The subjunctive mood in New World Romance varieties is frequently cited as an example of contact-induced erosion (see e.g., Garcia 1985; Gutiérrez 1994; Lantolf 1978; Ocampo 1990; Silva-Corvalán 1994b for Spanish and Laurier 1989 for Canadian French). We explore this phenomenon in more detail below. Other morphosyntactic elements said to be receding as a result of contact include the BE auxiliary (*sein, zijn, essere*) in Australian German, Dutch and Italian respectively, in favor of *haben, hebben, avere* 'have', on the model of English, which admits only *have* in such contexts (Clyne 2003: 118–119). Clyne does acknowledge regional variation in auxiliary usage, even in some non-contact varieties of these languages, but follows Thomason and Kaufman (1988: 58–59), in arguing that contact reinforces language-internal tendencies. And these examples could be multiplied *ad libitum*.

Much of the evidence for contact-induced change may lend itself equally well to alternate explanations. Sometimes inherent variability is equated with change (as in the aforementioned Australian example). Elsewhere, perceived differences from a standard are designated changes (as in the increase of overt subject pronouns). And it is often unclear whether the change in question is idiosyncratic or community-wide. The next section summarizes these and other pitfalls and details some of the considerations that must be taken into account in the empirical establishment of change.

3. Establishing change

3.1. Variability is not change

The hallmark of speech is inherent variability, alternate ways of saying the same thing (Labov 1969). Variability is a necessary condition for change, but is not, in and of itself, coterminous with it. In many areas of the grammar, alternations among variant forms may persist for centuries, but linguists who believe that language is invariant often interpret them as signs of change. This inference is particularly prevalent when the forms in question are detected among speakers or groups considered vulnerable to external influence (e.g., bilinguals, residents of minority-language communities in intense contact with a majority language, etc.). Change is also routinely invoked when one of the variants is salient or stigmatized, especially when a null form is involved. Thus, the widespread variability between subjunctive and indicative morphology in (prescribed) subjunctive-selecting contexts cited above is usually characterized in contact situations as *loss* of the subjunctive, despite the fact that this alternation has been attested for centuries (Poplack 2003; St-Amand 2002). So a first requirement is to determine whether variability is involved in change. Variant choice is never random, but is instead highly constrained by features of the (linguistic and extra-linguistic) environment. Measuring the trajectory of variant use over time, coupled with analysis of the fine conditioning of variant choice, provide a sensitive and refined diagnostic of change, in ways we illustrate in section 5 below.

3.2. Locating an earlier stage

At the most trivial level, change shows up as a *difference* (qualitative or quantitative) between an outcome and an earlier stage, but in much of the current literature, the structure of the earlier stage is either not appealed to, or at best, only inferred (e.g., Winford 2003: 64). Absent systematic comparison with a relevant precursor, any claims about what has changed must remain speculative. (Indeed, lack of an appropriate diachronic baseline is at least partially responsible for the widespread assumption that many salient features of contemporary vernaculars, both contact and non-contact, are recent innovations; see also Zwicky 2005).

The first step in establishing the existence of change is comparison over time. This may not be simple or straightforward, given the often fragmentary nature of surviving diachronic evidence. This abiding (and at times insoluble) problem in historical linguis-

tics has given rise to an entire methodology devoted to reconstruction (e.g., Campbell 1999, 2003; Hoenigswald 1960; Lehman 1992; Ringe 2003, to name but a few). Synchronically, however, with a modicum of ingenuity, earlier stages or surrogates for them can often be found.

3.3. The nature of the reference point

Not every earlier stage is equally useful in assessing change, however. In most reported cases, the comparison point (diachronic or synchronic) is some poorly defined "standard". Constituted only of forms that have been prescriptively ratified, the standard rarely acknowledges variants that cannot be associated (rightly or wrongly) with variant functions or readings (Poplack et al. 2002). As a result, the language ratified as standard may diverge considerably from the language actually spoken, which often develops independently of normative injunction (Poplack and Dion to appear; Leroux 2003; Poplack and Malvar 2007; Willis 2000). The variability most characteristic of vernacular speech involves competition among standard and nonstandard features. When compared with an idealized invariant standard, the latter may appear to be changes.

A nice illustration of the importance of an appropriate reference, or baseline, variety comes from a recent study of Spanish in contact with English in New York City (Otheguy, Zentella and Livert 2007). As noted above in connection with the extension of minor-use patterns, rates of overt pronouns are thought to be increasing among second-generation Spanish speakers, on the model of English, in which the subject is assumed to be categorically expressed (but see Leroux and Jarmasz [2006] for conditions governing variable subject pro-drop in English). Of course, *some* overt pronoun expression is tolerated in Spanish, but the standard baseline does not specify how much. What then constitutes an "increase"? Vis-à-vis what? Otheguy, Zentella and Livert's comparison with one surrogate pre-contact stage of Spanish (spoken by newcomers to New York) revealed, contrary to the widely espoused convergence hypothesis enunciated above, an even *higher* rate of overt pronouns than that displayed by their New York born-and-raised counterparts (2007: 786). Only once the appropriate comparison variety was identified (i. e. by distinguishing two newcomer dialects, each with a different *initial* rate of overt pronouns), were they able to demonstrate that pronoun expression had in fact increased in the second generation.

Invidious comparisons can only be avoided if the reference variety is commensurate with the variety hosting the candidate for change. If it is a spoken vernacular, the baseline should also be a spoken vernacular (or an appropriate surrogate thereof); if it is a contact variety, the baseline should be a pre-contact variety, and so on. As we will show below, when the inference of change is pursued scientifically via systematic comparisons (vertical and horizontal) of *appropriate* reference varieties, potential candidates often turn out either not to be changes, or not to be contact-induced, but rather cases of garden-variety inherent variability firmly rooted in the internal structure of language (Leroux and Jarmasz 2006; Poplack et al. 2006). Efforts invested in locating and mining appropriate earlier stages have paid off handsomely (Ayres-Bennett 2000; Gordon et al. 2004; Martineau and Mougeon 2003; Poplack and St-Amand 2007). The procedure is a good deal more laborious than simply inferring what the precursor was like, but in view of the potential for misidentification, we see this as a requirement for any claim of change based on synchronic data.

3.4. The social embedding of change

Another major impediment to detecting change resides in the nature, quantity and quality of the evidence. Manifestations of change are relatively rare in running discourse. This may be why much of the published evidence comes from data collected in experimental or quasi-experimental conditions (e.g., informant elicitation, subjective reaction tests, even "forced" interviews with semi-speakers [see, e.g., Altenberg 1991; Huffines 1991; Kaufman and Aronoff 1991; Montrul 2002; Schmidt 1991]). Atypical situations may entail atypical behavior, with little or no bearing on what transpires in real-life bilingual settings (Poplack 1983: 116). Without knowing how well, if at all, these experiments reflect language use in the community, their results cannot be interpreted.

This brings us to the distinction between *innovation* and *change*. Every change involves an innovation in the speech of one or more individuals and its subsequent diffusion across the rest of the community (Croft 2000; Joseph and Janda 2003: 13; Milroy 1993: 223; Shapiro 1995: 105, fn. 1; Thomason 2007). For some linguists, the innovation *is* the change, and it is assumed to occur during first language acquisition. Theoretically, we would argue (along with Labov 1994: 310–311 and most other sociolinguists), that "an innovation in a speaker's output is not a linguistic change until it has been agreed on and adopted by some community of speakers" (Milroy 1992: 221), *a fortiori* because the vast majority of changes do not "catch on" (e.g., Thomason 2001: 130). From an empirical perspective, however, this is a moot point: in all but the rarest of cases, the moment of innovation is simply unobservable.

We can rephrase the issue in terms of the relationship between the individual and the community. Claims of change are often based on exiguous speaker samples (e.g., Toribio's 2004 study of two individuals), if any. The linguistic behavior of the individual may be ephemeral, situation-bound and/or simply deviant from the norms of the remainder of the community (Labov 1972a; Poplack 1983: 119), especially when this behavior is elicited under laboratory conditions. Alternatively, it may be a predictable outcome of one or more explanatory factors, linguistic and/or extra-linguistic. Only once the speaker's linguistic behavior is situated with respect to that of the appropriate reference variety (*other* speakers of the same age, bilingual ability, educational level, etc.) is it possible to decide. If change has occurred, it should be observable (at least to some degree) in the speech of the relevant members of the community (e.g., the younger cohort, if spearheaded by youth; the highly bilingual, if induced by contact, etc.). In other words, in the absence of *diffusion*, change cannot be confirmed.

3.5. Type of bilingual population

The locus of the candidate for contact-induced change is also important. As Bullock and Toribio (2004: 91) observe, researchers have drawn on an array of bilingual populations, including heritage language speakers, incomplete acquirers and those undergoing language loss. Shifting populations have garnered most attention (see, e.g., Dorian 1981, 1989; Gal 1979; Mougeon and Beniak 1991; Otheguy, Zentella and Livert 2007; Silva-Corvalán 1991), but the linguistic behavior they display is not necessarily applicable to stable bilingual communities (Poplack 1993; Winford 2003: 64). In the widely studied

contexts of language restriction and obsolescence, structural differences (from the mainstream or non-contact variety) may in fact be inevitable, given that many of the informants may not (or no longer) use it with any regularity. In effect, then, although ethnically minority-group members, they may be L_2 speakers of the minority language (Poplack 1997: 305). L_2 speakers can expected to make all manner of acquisition errors, depending on their level of proficiency at the time of data collection. Though these may be transitory, in contact situations, they too are often construed as changes.

3.6. The linguistic embedding of change

In addition to a certain level of diffusion across the community, the innovation should also occur (to some degree) in relevant linguistic contexts. Relevant contexts are determined by situating the innovation in the larger host linguistic system—synchronic and diachronic. Among the questions to be answered are: To what extent has the innovation gained a foothold? What is its current role? Has it replaced a native form in one or more functions? Did it introduce a new function into recipient-language grammar? Isolated, anecdotal or exceptional examples, so often cited in the literature, reveal nothing about the regularity of the innovation, its productivity, or the extent to which it is entrenched in the language (Weinreich, Labov and Herzog 1968: 185). This can only come from studying the sustained discourse of speakers in context. Because changes examined synchronically are likely to still be in progress, and thus continue to feature variability, these questions can only be addressed by establishing the *structure* of the variability, as it emerges from the constraints conditioning variant choice.

3.7. Recognizing change

Most reported changes are tangible: a novel lexical element (e.g., a borrowing), unwarranted presence of an existing form (e.g., increased pronoun expression in pro-drop languages), absence of a prescribed form (e.g., the French subjunctive), or choice of the (prescriptively) "wrong" form (e.g., auxiliary *haben* for *sein* in Australian German or *avoir* for *être* in Canadian French). But observed changes tend to be far more subtle, especially during the sometimes very lengthy transition period from one language state to another. Linguists who focus on inherent variability recognize, in ascending order of importance, change in frequency of one or more competing forms, change in statistical significance of one or more factors contributing to variant choice and change in linguistic structure. With rare exceptions, these elude casual observation, but can be detected with the aid of the methods described in section 4.

3.8. Distinguishing contact-induced change from internal evolution

Once it has been established that a change has indeed taken place, it remains to demonstrate that it is contact-induced and not the product of drift. This step is usually skipped in the literature, much of which is predicated on the aforementioned (widespread but

unfounded) assumption that linguistic differences occurring in bilingual contexts are necessarily 1) changes and 2) contact-induced. Thomason (2007: 42) defines contact-induced change as follows: "a particular linguistic change is caused at least in part by language (or dialect) contact if it would have been less likely to occur outside a particular contact situation". Given how little we know about the *likelihood* of particular changes – within or outside contact situations – we offer alternative criteria:

> A candidate for contact-induced change in a contact variety is present in the presumed source variety and either 1) absent in the pre-contact or non-contact variety, or 2) if present (e.g., through interlingual coincidence), is not conditioned in the same way as in the source, *and* 3) can also be shown to parallel in some non-trivial way the behavior of a counterpart feature in the source.

Compliance with these conditions can only be ascertained through systematic quantitative comparisons, of a *diagnostic* linguistic feature, with an earlier or pre-contact stage (or a surrogate thereof), with a non-contact variety, and most important, with the presumed source. As we shall see in what follows, these criteria have not been met by most extant studies, with the result that the existence, frequency and extent of contact-induced change remain unclear.

Summarizing, sufficient understanding of the outcomes of contact appears beyond the reach of anything but systematic corpus-based research carried out within the bilingual community. Application of an empirically accountable quantitative methodology to large and representative bodies of bilingual speech can contribute significantly to resolving the difficulties associated with ascertaining whether contact-induced change has authentically occurred. It is to such a methodological framework that we now turn.

4. Applying the comparative variationist framework to the investigation of change

4.1. The variationist paradigm

The approach that we exemplify in the remainder of this article draws on the combined resources of variationist sociolinguistics (e.g., Labov 1972a, Labov 1972b; Labov et al. 1968; Weinreich, Labov and Herzog 1968) and the Comparative Method of historical linguistics (e.g., Baldi 1990; Campbell 1998; Hoenigswald 1960; Meillet 1967). A central tenet of the variationist paradigm is that the recurrent choices made by speakers in discourse, though not completely predictable, are not random, but are subject to constraints imposed by "the phonological environment, the syntactic context, the discursive function of the utterance, topic, style, situation and personal and/or sociodemographic characteristics of the speaker or other participants" (Sankoff 1988a: 151). The key theoretical construct of this framework is the *linguistic variable* (Labov 1966/1982), which comprises a set of *variants* among which speakers alternate in expressing a given meaning or function. This results in a heterogeneous but structured linguistic system. A foundational working principle is that the structure – grammatical and social – underlying the heterogeneity can be inferred from the distribution and conditioning of competing variants in discourse.

4.2. Communities and speakers

We noted above that change proceeds from diffusion, and extent of diffusion can only be assessed in reference to a community of speakers. A crucial step, though often neglected, is to locate an appropriate (here, bilingual) *community* in which to carry out the study. The nature and makeup of the community allow us to test hypotheses about the extra-linguistic factors often claimed to be relevant to contact-induced change. These include *intensity of contact* (the more intense, the greater the likelihood of structural interference), *length of contact* (the greater the time-depth, the greater the likelihood of change), *status* of the languages in the community (minority or majority) and *size* of speaker population (the fewer and more marginal the speakers, the more likely that they will borrow from the language of the dominant group) (e.g., Romaine 1995; Thomason 2001; Winford 2003). Study sites can be selected to represent these parameters and to test whether these factors are or are not operating.

Speakers constituting these communities can be sampled to represent individual dimensions of the contact axis. Sampling protocols are also guided by hypotheses about how contact-induced change works. According to one, individual level of bilingualism is inversely correlated with mastery of the minority-language grammar, such that the greater the proficiency in the other language, the more likely the possibility of simplification or loss of minority-language linguistic structure. Another proposes that the incorporation on the local level of majority-language material (via code-switching or borrowing) into otherwise minority-language discourse may bring with it associated grammatical properties, and these in turn may lead to structural convergence (e.g., Backus 2005; King 2005; Muysken 2000; Thomason 2001; Winford 2003). According to this hypothesis, a speaker's individual propensity to code-switch would be a predictor of contact-induced change. Individual attitudes toward, and relative prestige of, each of the languages could also affect the extent and direction of change (e.g., Appel and Muysken 1987: 158; Joseph 2007; Romaine 1995: 66; Thomason 2001).

The case studies in section 5 demonstrate how these and other hypotheses can be tested on linguistic production data provided by speakers representing different combinations of these conditions, thus enabling us to ascertain which are most explanatory of contact-induced change and which community members are most likely to lead it.

4.3. Sampling bilingual speech

Once the appropriate speakers have been identified, we can record samples of their spontaneous speech (bilingual *and* monolingual) in quantities sufficient to permit accountable analysis. The speech of choice for such an endeavor is the *vernacular*. Not only does the vernacular offer "the most systematic data for linguistic analysis" (Labov 1984: 29), it is also the register in which linguistic manifestations of language contact are most likely to occur. Variationists mine vernacular corpora to uncover the regular patterns that characterize natural exchanges in the speech community and to distinguish them from isolated or idiosyncratic occurrences. A pattern may be defined as "a series of parallel occurrences (established according to structural and/or functional criteria) occur-

ring at a non-negligible rate of use" (Poplack and Meechan 1998: 129). Patterns can only be discerned by systematic quantitative analysis of the data in accordance with the principle of *accountability* (Labov 1972b: 72) . This requires not only that all the relevant occurrences of a feature be incorporated into the analysis, but additionally all the contexts in which it could have occurred but did not. Adherence to the principle of accountability allows the analyst to identify anomalous occurrences which are neither representative nor recurrent (and thus unlikely to trigger change), while identifying regular tendencies, which *may* (or may not) herald change.

4.4. Variable rule analysis

Certain linguistic and extra-linguistic conditions can be expected to promote change, though none is fully determinant. Claims that even favorable linguistic factors can be overridden by the right constellation of social factors (e.g., Thomason 2001: 77; Winford 2003: 25) raise the issue of the relative weights each contributes to a particular outcome and the relationship of such weights to different types of contact situations. This is a problem that lends itself well to multivariate analysis, the major analytical tool in the variationist framework. *Variable rule analysis* (Rand and Sankoff 1990; Sankoff, Tagliamonte and Smith 2005) facilitates extraction of regularities from the welter of data making up running discourse. The multiple regression procedure incorporated within it helps determine which aspects of the linguistic (e.g., phonological, morphological, syntactic, etc.) and extra-linguistic (e.g., minority status, bilingual ability, age, etc.) contexts contribute statistically significant effects to variant choice when all are considered simultaneously (e.g., Sankoff 1988b). These aspects are *hypotheses* about the structure of variant choice. Variable rule analysis enables us to operationalize them (as factors) and determine which give the best account of the data. The goals are to ascertain whether change has taken place and if so, whether it is the product of contact or internal development.

Much empirical work on contact-induced change in progress revolves around frequencies. As detailed by Poplack and Tagliamonte (2001: 92), however, differences in rates of variant selection must be used with caution to infer change, contact-induced or otherwise. The key diagnostic in assessing the relationship and provenance of forms is the *constraint hierarchy*, the configuration of environmental factors affecting the probability that a given variant form will be selected, along with the direction of their effects. We construe these as a portion of the grammar underlying the variability. To the extent that constraint hierarchies are language-specific, they are powerful indicators of change (see section 5).

Similarities in surface form in contact and source varieties can be misleading. They may result from borrowing or transfer, justifying the inference of change, but may also be due to interlingual coincidence or to linguistic universals. To rule out alternative explanations, we rely on *conflict sites* (Poplack and Meechan 1998: 132): functional, structural and/or quantitative *differences*, typically manifested as a conflict in constraint hierarchies. The conflict site, together with detailed cross-variety comparison, plays a crucial role in detecting change and identifying its ultimate provenance. If the hierarchy of constraints conditioning the variable occurrence of a candidate for change (e.g., Ø subjunctive, overt subject pronoun, preposition stranding, null relative marker, periphrastic future) in a contact variety is the same as that of its pre-contact precursor, while

differing from that of its presumed source, no structural change has taken place. If it features a constraint hierarchy different from those of both its pre-contact precursor *and* the presumed source, we can infer that change has occurred, but not one that is contact-induced. Only when a candidate for change in a contact variety features a constraint hierarchy *different* from that of its pre-contact precursor, but *parallel* to that of its presumed source, can we conclude in favor of contact-induced change.

4.5. Cross-variety comparison

Comparisons are at the heart of inferring change, and as such, figure, at least implicitly, in all studies of contact-induced change. But many have been carried out without reference to the principle of accountability, or involve an idealized comparison variety. Still others are simply anecdotal. Innovative in the research program we describe here is the extent of the scientific apparatus brought to bear on the comparison. The comparative variationist enterprise (Poplack and Meechan 1998; Poplack and Tagliamonte 2001; Tagliamonte 2002) contributes not only a methodological aspect – the construction, statistical analysis and linguistic interpretation of suitable comparison varieties (pre-contact, non-contact and post-contact), but also the critical capacity to decide among competing hypotheses. The focus is on identifying empirical criteria capable of both detecting change and ascertaining its source, and testing hypotheses empirically to determine the goodness of fit with the data. This involves principled data collection, enunciation and operationalization of hypotheses and their statistical evaluation in large-scale corpora. In section 5, we illustrate this approach by considering a number of case studies which attest to the versatility of quantitative variationist methodology for identifying contact-induced change.

5. Putting contact-induced change to the test

The studies we present below were carried out in a context of long-term intense contact between French and English in Canada, a situation that meets all the criteria considered favorable to structural convergence (e.g., Thomason and Kaufmann 1988; Winford 2003: 90). Depending on locality, each of the languages has a different official status: English is the majority language in Ontario, while in Quebec, French assumes this role. This configuration allows for comparison of each of the languages in minority and majority guise. The 284 speakers involved in these studies are distributed among these four conditions. Other characteristics relevant to the contact situation include density of contact on the local level (as expressed by the ratio of anglophones to francophones in the community), individual bilingual ability (as measured by a bilingual proficiency index [Poplack 1989, Poplack, Walker and Malcolmson 2006]) and relative propensity to borrow or code-switch (as inferred by comparing individual rates of each across sample members). Standard extra-linguistic factors like social class, education, etc. were also taken into account. The inference of change over time was investigated by comparing the speech of younger and older generations.

The linguistic data come from two massive corpora of spontaneous speech collected from these speakers: the *Ottawa-Hull French Corpus* (Poplack 1989) and the *Quebec English Corpus* (Poplack, Walker and Malcolmson 2006). The typological similarity of French and English should render them particularly hospitable to contact-induced change. In addition, superficially similar structures were deliberately targeted, as these are considered to be prime candidates for grammatical convergence. Indeed, at least two of the cases described below had been reported to result from this process.

5.1. Attrition of the subjunctive in Ottawa-Hull French

In the most straightforward case, we locate, in community speech, a linguistic feature that is used differently from the way it is prescribed and determine which individuals are deviating the most. If the difference is the result of change, and if that change is contact-induced, its leaders should figure among the younger, the more bilingual and those in most intense contact with the majority language. The French subjunctive is an excellent candidate. It is (prescriptively) obligatory after the class of "subjunctive-selecting" main verbs, like *falloir* in example (1), but in Canadian French it has been found to be variable ((1); Poplack 1990). (This and all ensuing examples are reproduced verbatim from audio recordings from one of the corpora described above. The codes in parentheses refer to corpus [QEC: *Corpus of Quebec English*; OH: *Corpus of Ottawa-Hull French*], speaker and line number. The examples cited throughout this paper are instantiations of *patterns* that emerged from the large-scale quantitative analyses presented in sections 5.1 – 5.3)

(1) *Fallait* *qu' elle* ***répond*** (IND) *"oui, tu peux faire trois pas de*
Necessary (PAST) that she answer (IND) yes you can make three steps of
géant".
giants
Fallait *qu' elle* ***réponde*** (SUBJ) *la phrase complète.*
Necessary (PAST) that she answer (SUBJ) the sentence complete.
'She had to say "yes, you may take three giant steps". She had to say the whole sentence.' (OH.025.2186)

Because the subjunctive is no longer used productively in English, the (variable) occurrence of the indicative in contexts like (1) is inferred to be a stage in the loss of this form, a development which Laurier (1989) characterized as the result of contact.

Table 22.1 displays the results of a multivariate analysis of the extralinguistic factors contributing to choice of subjunctive. In this and the ensuing tables, the higher the figure, the greater the probability the "application" variant will be selected in the environment listed on the left. Table 22.1 reveals a small, but statistically significant, tendency for the most bilingual speakers to use least subjunctive. This result supports the hypothesis that high levels of bilingualism lead to linguistic simplification and loss. Most studies would stop at this demonstration (if they provided any quantitative information at all). But if this is a contact-induced change, effects of the other factors should point in the same direction. This not the case. Instead, the two French-majority-language communities in Québec, Vieux Hull and Mont Bleu, which should favor the subjunctive most under this

Tab. 22.1: Variable rule analysis of the contribution of extralinguistic factors to the choice of subjunctive under verbal matrices

Corrected mean		*0.715*
Total N		2694
SEX		
Female		.52
Male		.46
AGE		
45–54		.55
55–64		.53
15–24		.52
25–34		.50
65+		.48
35–44		.44
NEIGHBOURHOOD OF RESIDENCE		
Vieux Hull, Quebec	(FR majority working class)	.57
West End, Ontario	(FR minority working class)	.51
Vanier, Ontario	(FR minority upper-working class)	.50
Basse-Ville, Ontario	(FR minority working class)	.50
Mont Bleu, Quebec	(FR majority upper-middle class)	.43
ENGLISH PROFICIENCY		
mid-low		.54
low		.53
mid-high		.48
high		.42

hypothesis, instead use most *and* least, depending on social class. (Working-class speakers, even more counter-intuitively, apparently use it the most.) Also militating against the inference of change is the fact that the age distribution is not monotonic: the youngest speakers are as likely to select the subjunctive as their oldest counterparts. Further analysis revealed that these apparent extra-linguistic effects are in fact epiphenomena of a strong lexical effect (Poplack 1990, 1992: 2001). Subjunctive morphology is highly associated with only a few main verbs, and because those verbs happened to be unevenly distributed across the subgroups represented in Table 22.1, this showed up as a discrep-

Tab. 22.2: Variable rule analysis of the contribution of *English proficiency* to the choice of subjunctive in verbs embedded under four classes of verbal matrix

	English Proficiency			
	low	mid-low	mid-high	high
VERBS EMBEDDED UNDER				
high frequency/high subjunctive matrix verb: *falloir*	.55	.58	.40	.42
high frequency/high subjunctive matrix verbs: *vouloir / aimer*	.46	.58	.42	.60
low frequency/low subjunctive matrix verbs	.68	.47	.49	.31
low frequency/variable subjunctive matrix verbs	.34	.60	.61	.39

Tab. 22.3: Variable rule analysis of the contribution of *socioeconomic class* to the choice of subjunctive in verbs embedded under four classes of verbal matrix

	Socioeconomic class			
	unskilled	skilled	sales / service	professional
VERBS EMBEDDED UNDER				
high frequency/high subjunctive matrix verb: *falloir*	.45	.53	.47	.69
high frequency/high subjunctive matrix verbs: *vouloir / aimer*	.36	.62	.64	.52
low frequency/low subjunctive matrix verbs	.56	.47	.42	.58
low frequency/variable subjunctive matrix verbs	.50	.42	.42	.68

ancy in subjunctive usage. When the lexical identity of the main verb is factored out, as in Table 22.2, bilingual proficiency no longer shows a systematic effect.

The real explanation for differential use of the subjunctive is class-based, and is not relevant to the contact situation at all, as illustrated in Table 22.3 (see Poplack 1997 for detail). This result could not have been intuited from casual comparisons with either the idealized standard benchmark (which prescribes the subjunctive categorically), or amongst proficiency groups. This is but one illustration of our earlier contention that differences in overall rates of variant occurrence need not be indicative of change, contact-induced or otherwise. Indeed, real-time analysis over a time span of 120 years (St-Amand and Poplack 2002) showed that the current situation of the subjunctive is actually a retention.

5.2. Preposition stranding in Ottawa-Hull French

It has been argued that a local factor, the extent to which individual speakers draw on the two languages in discourse, is a better predictor of contact-induced change than macro-level factors like minority status or density of contact. Under this hypothesis, speakers who code-switch the most would lead contact-induced change (Backus 2005; Bullock and Toribio 2004; King 2005; Thomason 2001; Winford 2003). Zentz (2006) was the first to test this hypothesis empirically, using as a diagnostic the syntactic phenomenon of preposition stranding in Canadian French, as in (2).

(2) *Comme le gars que je sors* ***avec****, lui il parle- bien il est français*
Like the guy that I go out with him he speaks- well he is French
'Like the guy I'm going out ***with***, he speaks- well he *is* French.' (OH.013.2122)

Because preposition stranding is normatively proscribed, examples like (2) have been ascribed to contact, if not convergence, with English (e.g., King 2000; Roberge and Rosen 1999), in which stranding is presumed to occur freely.

But the phenomenon of stranding cannot be fully understood in isolation from the behavior of the remainder of the preposition system. In particular, French does admit phrase-final prepositions in another, superficially similar, construction, known as *orphaning* (Zribi-Hertz 1984). This is illustrated in (3). In contrast to stranding, orphaning is not admissible in English, as can be seen from the gloss of (3).

(3) *Si tu veux l' avoir avant, il faut tu payes* ***pour.***
If you want it to.have before it have you pay for
*'If you want it before, you have to pay *for*.' (OH.013.258)

This raises the question of whether preposition stranding in French is an extension (to relative clauses) of the native orphaning, or a change induced by contact with English. Reasoning that if the latter were the case, and the above claims were correct, speakers who code-switched frequently to that language should engage in more stranding, Zentz (2006) contrasted the stranding behavior of "high" and "low" code-switchers.

Using constraint hierarchies to first determine and then compare the structure of the choice mechanism, Zentz showed (Table 22.4) that the strongest predictor of preposition orphaning in French is lexical.

"Strong" prepositions (a category made up, for the most part, of four: *avec* 'with', *pour* 'for', *dedans* 'in', *dessus* 'on') favor orphaning, while "weak" prepositions (here, *à* 'to' and *de* 'from') disfavor it almost categorically. This, pattern, which is perfectly consistent with normative injunctions (Grevisse 1980: 1509), is shared by all speakers, regardless of propensity to code-switch, as can be seen by the parallel constraint hierarchies. In comparison, stranding, the candidate for change, not only occurs at the same low rate as orphaning (approximately ten percent of all eligible contexts), but is constrained in the same way.

Here too, the most important conditioning factor is the lexical identity of the preposition. As with orphaning, stranding is also eschewed with weak prepositions and highly promoted by strong prepositions. And the strong prepositions in question are precisely those that favor orphaning. This effect is again as true of high as of low code-switchers.

Once stranding – the superficially English structure – is examined in the context of the recipient-language system, three independent results counter the notion that it was borrowed from English. Stranding occurs at the same (low) rate and more important, obeys the same variable constraints as native orphaning. Like orphaning, stranding is restricted to like orphaning, standing is basically restuded to the same four strong prepositions. In English, on the other hand, stranding would seem to be equally productive with a wide variety of lexical prepositions, strong and weak. Finally, there is no distinction between high and low code-switchers, refuting claims that the former are agents of contact-induced change (see also Torres-Cacoullos and Travis to appear).

The great merit of the Zentz study is that it marshals accountable scientific methodology to demonstrate that, contrary to appearances, a syntactic structure apparently identical to a counterpart in the majority language, and which is moreover (prescriptively) proscribed in the minority language, is nonetheless unlikely to be a contact-induced change. This is because the conditions giving rise to stranding in French parallel those constraining native orphaning, while *differing* from those assumed to be operating in English. The drawback of this study (and most others) is that it compares the situation

Tab. 22.4: Independent variable rule analyses of the contribution of linguistic factors to preposition orphaning and stranding for low and high code-switchers (brackets indicate that the factor was not statistically significant)

	Orphaning		Stranding	
	"Low" Code-switchers	"High" Code-switchers	"Low" Code-switchers	"High" Code-switchers
Corrected mean	*.09*	*.02*	*.01*	*.03*
Total N	982	662	224	116
SYNTACTIC FACTORS				
Place of topic				
Discourse referent	[ ]	[ ]	N/A	N/A
Intra-sentential referent	[ ]	[ ]	N/A	N/A
Type of construction				
Cleft	N/A	N/A	.35	[ ]
Object precedes relative	N/A	N/A	.40	[ ]
Pseudo-cleft	N/A	N/A	.99	[ ]
Presence of another complement				
None	[ ]	[ ]	[ ]	[ ]
Additional Post-V complement	[ ]	[ ]	[ ]	[ ]
Clitic	[ ]	[ ]	[ ]	[ ]
SEMANTIC FACTORS				
Semantic content of preposition				
"Weak"	0	.19	.19	.19
"Strong"		.96	.99	.95
Obligatoriness of preposition				
Obligatory	.42	[ ]	[ ]	[ ]
Non-obligatory	.61	[ ]	[ ]	[ ]
Humanness of noun complement				
Non-human	.57	[ ]	.46	[ ]
Human	.36	[ ]	.82	[ ]

in the contact language with what is *presumed* to be the case in the source. Absent any accountable study of preposition stranding in the contact variety of English, we cannot be fully confident about exactly *what* was transferred. In additional, there is no evidence (in this or other studies), that the current state of affairs represents a *change*. Demonstration of contact-induced change would minimally include a vertical comparison (preferably with a pre-contact stage, but at the very least with an earlier stage), a horizontal comparison with a non-contact variety and comparison with the structure of the source.

5.3. Relative marker expression in Quebec English

A study by Poplack et al. (2006) illustrates the unexpected results to emerge from such an approach. It involves the variable use of overt and null relative markers in accusative relative clauses in Quebec English, a minority language in intense contact with French. This is illustrated in example (4). Quebec French shows the same (surface) alternation in the same context (5).

(4) a. *I said "Oh yeah, I remember I was wearing this disgustingly ugly green suit* ***that*** *you made me wear", or something.* (QEC.031.1135)

b. *I was always known as Carrot Head and Freckle Face and a few other very vulgar names* ***Ø*** *they called me.* (QEC.006.602)

(5) a. *Bien je comprenais le français, mais c' est les mots* ***que*** *je comprenais pas.*
Well I understood the French but it is the words that I understood not
'Well, I understood the French, but it's the words *that* I didn't understand.' (OH.059.812)

b. *Les maisons* ***Ø*** *ils faisaient dans ce temps là le bois était très- il était pas cher pantoute.*
The houses Ø they made in that time there the wood was very- it was not expensive at all
'The houses they made in those days, the wood was very- it wasn't expensive at all.' (OH.050.698)

In particular, both languages permit a null variant (4b, 5b), although this is prescriptively unacceptable in French. How is it generated? Table 22.5 displays the factors constraining choice of null variant across mainstream varieties and cohorts of contact-variety speakers.

A first important observation is that the structure of variable relative marker expression differs in the two non-contact varieties. In English (column 2), the strongest predictor of a null relative marker is the presence of intervening material between the subject of the relative clause and its antecedent; in French (column 1), the major effect is phonological. Because these effects are language-specific, we can use them as diagnostics of contact-induced change. We may now compare the contact variety to its non-contact counterpart (columns 2 and 3).

For all factor groups, the constraint hierarchies are parallel. This means that, in the aggregate, the contact variety of English is no different from its mainstream counterpart. But according to the hypotheses enunciated earlier, certain subgroups of the population could be expected to spearhead changes. So we will want to establish whether *all* of the minority-language speakers behave in the same fashion, by comparing the conditioning of variant choice according to speaker age and bilingual ability. Predictably, the behavior of the older (column 5) and least bilingual (column 7) speakers is very similar to that of speakers of noncontact English (column 2) in terms of constraint hierarchies. But the younger (column 4) and more bilingual speakers (column 6) share less with the main-

Tab. 22.5: Independent variable rule analyses of the contribution of linguistic factors to the selection of the *null* relative marker: across mainstream varieties

	1	2	3	4	5	6	7
	Mainstream French	Mainstream English	Contact English	Younger speakers	Older speakers	Most bilingual speakers	Least bilingual speakers
Corrected mean:	*.019*	*.367*	*.532*	*.400*	*.502*	*.503*	*.568*
Total N	50/460	109/288	150/281	131/316	128/253	75/146	75/135
Intervening material							
None	[ ]	.52	.54	.55	[ ]	.57	[ ]
Present	[ ]	.17	.24	.19	[ ]	.14	[ ]
Antecedent type							
Pronoun	.71	.73	[ ]	[ ]	.76	[ ]	.89
Noun	.44	.46	[ ]	[ ]	.45	[ ]	.42
Type of subject of the relative clause							
Pronoun		.52	.51	[ ]	.53	[ ]	.53
Noun	0	.28	.32	[ ]	.15	[ ]	.18
Following phonological environment							
Sonorants	.13	[ ]	[ ]	.54	[ ]	[ ]	[ ]
Obstruents	.94	[ ]	[ ]	.36	[ ]	[ ]	[ ]
Definiteness of antecedent							
Definite	[ ]	[ ]	[ ]	[ ]	.59	[ ]	.60
Indefinite	[ ]	[ ]	[ ]	[ ]	.36	[ ]	.31
Sentence type							
Cleft	.61	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Other	.44	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

stream in terms of the fine conditioning of variant choice. This kind of result is consistent with their predicted role as leaders of change. Are these in fact changes? Comparison with older speakers suggests that they are. But are they contact-induced? As we noted earlier, to sustain any claim in this respect, the involvement of the source language (here, French) must be established. The three major "departures" of the younger and more bilingual cohorts with respect to their older and less bilingual counterparts involve the effects of antecedent type, type of subject of the relative and following phonological environment. The first two effects have been neutralized in the younger generation, while in French, as shown in column 1, they are strong and significant. The effect most relevant to the contact situation is phonological. Recall that this is the major predictor of variant choice in French, and the younger and most bilingual anglophones are the only ones who display it. This result is consistent with the inference of contact-induced change. But comparison of constraint hierarchies shows that the French effect goes in the *opposite* direction. (The apparent English phonological effect is an epiphenomenon of a developing tendency for certain pronominal subjects which are coincidentally obstruent-initial (*he*, *she*, *they*) to disfavor deletion.) Here again, differences among minority-language cohorts, suggestive as they appear on the surface, cannot be attributed to convergence

Tab. 22.6: Independent variable rule analyses of the contribution of linguistic factors to the selection of the *null* relative marker: younger contact vs. benchmark speakers

	Younger contact speakers	Younger mainstream speakers
Corrected mean:	*.482*	*.299*
Total N	83/168	48/148
Intervening material		
None	.56	.55
Present	.18	.10
Antecedent type		
Pronoun	[ ]	.76
Noun	[ ]	.45
Type of subject of the relative clause		
Pronoun	[ ]	[ ]
Noun	[ ]	[ ]
Following phonological environment		
Sonorants	[.53]	[.54]
Obstruents	[.38]	[.34]
Definiteness of antecedent		
Definite	.59	[ ]
Indefinite	.40	[ ]

with French. To what then are they due? Comparison with the non-contact benchmark provides the answer (Table 22.6). When speakers of mainstream English are analyzed according to age, we find the same apparent changes in progress.

In this case our method has enabled us to detect change – the young minority-language anglophones are diverging from their elders (column 5 of Table 22.5) – and to attribute it to the correct source: participation in a change in progress in mainstream English.

These examples, along with many others (e.g., Dion 2003; Jarmasz 2005; Leroux and Jarmasz 2006; Petrik 2005; Poplack and St-Amand 2007; St-Amand 2002) demonstrate that when the inference of contact-induced change is pursued systematically, it becomes increasingly difficult to justify. Particularly instructive is the fact that even in quantitative studies, differences between cohorts or overall rates may be masking effects that are independent of the contact situation.

6. Discussion

Change is widely considered a predictable, if not inevitable, outcome of language contact. The literature is rife with accounts of radical alterations of the grammars of contact languages, whether by virtue of incorporating forms or functions with no antecedent in the recipient language, or losing one or more of its core grammatical elements. Many of these changes would appear to be very rapid, occurring even in the space of a single

generation. This is consistent with another widespread, if unproven, idea: that contact *accelerates* change. On both of these counts, contact-induced change appears suspiciously different from internal evolution, which we know to be gradual, moderate and vastly more conservative.

Observing that scientific *demonstration* of change, contact-induced or other, was often missing from this research, we reviewed some of the phenomena that could lead the analyst down the garden path of inferring change, even where such an inference is not warranted. Chief among them are the inherent variability characteristic of spontaneous speech and the dearth of accountable corpus-based studies of the languages in contact. When confronted with an idealized standard variety, but lacking systematic comparison with the more relevant pre- and non-contact varieties, or with the supposed source variety of the claimed change, we detailed how the (variable) occurrence of non-standard variants could easily be confused with change.

We also outlined a number of conditions necessary to establish the existence of contact-induced change. Many had already been enunciated elsewhere (Poplack 1993; Poplack and Meechan 1998); some were recently reiterated (e.g., Backus 2005: 310). The following are paraphrased from Thomason (2001: 93–94):

1. Situate the proposed change with respect to its host linguistic system
2. Identify a presumed source of the change
3. Locate structural features shared by the source and recipient languages
4. Prove that the proposed interference features were not present in the pre-contact variety
5. Prove that the proposed interference features were present in the source variety prior to contact
6. Rule out (or situate) internal motivations

Thomason herself concedes that "in many, possibly even most contact situations around the world we cannot at present satisfy requirements 4 and 5", going so far as to caution that "if we can't satisfy all the requirements, then we can't make a solid case for contact-induced change" (Thomason 2001: 94). This raises the issue, which has preoccupied us throughout this article, of just how many of the contact-induced changes currently treated as facts satisfy *any* of these conditions, let alone all. This is a very important question, since these "facts" are the basis for our theories about the nature, extent and mechanisms of contact-induced change.

We observed earlier that for many diachronic changes, the crucial evidence would now be impossible to reconstruct. But as we have been at pains to illustrate, this is certainly not the case for synchronic situations. It is thus all the more puzzling that the requisite information is missing from all but a handful of contemporary studies. Why should this be? The growing, but patently false, perception among linguists working in other paradigms that it is too difficult, if not impossible, to access the relevant evidence may play a role. The treatment of frequency is a case in point. Frequency is at the root of most inferences of change, whether manifested as raw rates of occurrence of a form, or its diffusion across a community or linguistic system. Among the few who have addressed this issue at all, a certain skepticism about the feasibility of ascertaining frequency empirically is detectible. Thus with reference to the appearance of foreign elements in bilingual speakers' discourse, Thomason observes, "determining the frequency of appearance is a hopeless task" (2001: 134), and goes on to query: "how can one

determine whether a speaker used [an element] once, occasionally or frequently?" Likewise, while Backus (2005) admits that most changes are likely of the "frequency-changing kind", he contends that these are precisely the changes that "are virtually impossible to demonstrate empirically" (2005: 334). Such assertions betray a surprising lack of familiarity with corpus-based studies, and fly in the face of the important results of large-scale quantitative analyses that have been performed on languages in contact.

Despite recent concerns about the absence of an empirically well-founded framework for investigating bilingual speech (Van Coetsem 2000: 39–40; Winford 2007: 22) and the alleged limitations of variation-based techniques for illuminating contact phenomena (Muysken 2000: 250), we submit that the rigorous comparative methodology presented here is capable of detecting change and distinguishing external from internal motivations. It thus discriminates among competing hypotheses about the nature of these processes.

Agreeing with Thomason (2001: 88) that the crucial similarity between the processes of contact-induced and internal change has to do with *competition* (i.e. variability) between an innovative feature and an older feature, we have outlined a procedure for assessing contact-induced change within a framework that takes account of its symbiotic relationship with linguistic variation. We hope to have shown that reliance on the structure of variant choice affords a depth of analysis that transcends impressionistic observations based on perceived similarities between surface forms and rates of occurrence.

Indeed, the very strength of the variation-based approach resides in its capacity to discern even minor perturbations in constraint hierarchies, which *may* be harbingers of change, contact-induced or otherwise. Otheguy, Livert and Zentella (2007) provide the most recent confirmation that it can detect even the smallest adjustments in constraint ranking across generations of bilingual speakers.

Illustrating with studies of three well-documented candidates, we found no evidence of change in the first two, and the change documented in the third was not contact-induced. Moreover, in line with Thomason's (2001: 91–92) caveat that a "case for contact origin can't be made on the basis of single feature", additional research situating a wide variety of Canadian French and English vernacular variants in historical, social and linguistic context corroborates the findings presented here: none of the contemporary non-standard variants typically ascribed to either linguistic change or language contact are in fact innovations (Poplack and Dion to appear; Elsig and Poplack 2006; Jarmasz 2005; Leroux 2007; Poplack 1993). Rather, virtually all were not only present in the pre-contact variety, but, more important, they were largely constrained in the same way.

These results are especially striking when we consider that all the popularly invoked prerequisites for contact-induced change – social (protracted contact, extensive bilingualism; high levels of proficiency in the respective source languages, etc.) and linguistic (i.e. typological congruence [Thomason 2001: 71], interlingual equivalence [Backus 2005: 326]; grammatical "weak points" [Bullock and Toribio 2004: 91], attractiveness [Johanson 2002: 2]) – are met.

Why then did contact-induced change fail to occur? Given the present state of our knowledge, we can only speculate. The length of contact may have been too brief to result in change, although a time span of two centuries is a good deal longer than those on which many recent claims have been based. Perhaps the morpho-syntactic features we examined are particularly resistant to change, although they were specifically selected on the basis of typological similarities thought to facilitate it. We cannot rule out the possibility that our analytical tools may have misidentified or excluded changes; we

reserve judgment until a demonstration to that effect is offered. Certainly the circumstances – linguistic, social and attitudinal – could not have been more propitious. But while favorable circumstances may encourage change, we know that they do not determine it. A remaining possibility is that contact-induced change is a good deal less common than the literature suggests. This is our best guess at the moment.

Of course, the analyses we have presented, despite being based on corpora of bilingual speech with sizes ranging into the millions of words and large representative samples of individuals in contact situations, cannot begin to do justice to the full trajectory of change over time and across all speakers of a language. But opting for the alternative – anecdotal observation – simply because "it is much easier to show a qualitative change" (Backus 2005: 334) – ignores the important scientific advances in empirical linguistics made over the last half century. Contact-induced change is *not* an inevitable, nor possibly even a common, outcome of language contact. Only more accountable analyses of more contact situations will tell. In the interim, the burden of proof is on those who claim that it has occurred.

7. Acknowledgement

The research reported here was generously funded by grants to Poplack from the Social Sciences and Humanities Research Council of Canada and the Killam Foundation. Poplack holds the Canada Research Chair in Linguistics at the University of Ottawa.

8. References

Aikhenvald, Alexandra
2002 *Language contact in Amazonia.* New York: Oxford University Press.

Altenberg, Evelyn P.
1991 Assessing first language vulnerability to attrition. In: Seliger and Vago (eds.), 189–206.

Appel, René and Pieter Muysken
1987 *Language Contact and Bilingualism.* London: Edward Arnold.

Ayres-Bennett, Wendy
2000 Voices from the past: Sources of spoken seventeenth-century French. *Romanische Forschungen* 112(3): 328–348.

Backus, Ad
2005 Codeswitching and language change: One thing leads to another? *International Journal of Bilingualism* 9(3/4): 307–340.

Baldi, Philip
1990 Introduction: The Comparative Method. In: Philip Baldi (ed.), *Linguistic Change and Reconstruction Methodology*, 1–13. Berlin: Mouton de Gruyter.

Bakker, Peter
1997 *"A Language of Our Own": The Genesis of Michif, the Mixed Cree-French Language of the Canadian Métis.* (Oxford Studies in Anthropological Linguistics 10.) Oxford: Oxford University Press.

Bakker, Peter and Robert A. Papen
1997 Michif: A mixed language based on Cree and French. In: Sarah G. Thomason (ed.), *Contact Languages*, 295–363. Amsterdam: John Benjamins.

Bullock, Barbara E. and Almeida Jacqueline Toribio
2004 Introduction: Convergence as an emergent property in bilingual speech. *Bilingualism: Language and Cognition* 7(2): 91–94.
Bynon, Theodora
1977 *Historical Linguistics*. Cambridge: Cambridge University Press.
Campbell, Lyle
1998 *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
Campbell, Lyle
1999 *Historical Linguistics: An Introduction*. Cambridge, MA: MIT Press.
Campbell, Lyle
2003 How to show languages are related: Methods for distant genetic relationship. In: Joseph and Janda (eds.), 262–282.
Clyne, Michael
2003 *Dynamics of Language Contact*. Cambridge: Cambridge University Press.
Croft, William
2000 *Explaining Language Change: An Evolutionary Approach*. Harlow/New York: Longman.
Dawkins, Richard MacGillivray
1916 *Modern Greek in Asia Minor*. Cambridge: Cambridge University Press.
Dion, Nathalie
2003 L'effacement du *que* en français canadien: Une étude en temps réel. Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
Dorian, Nancy C.
1981 *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia, PA: University of Pennsylvania Press.
Dorian, Nancy C. (ed.)
1989 *Investigating Obsolescence: Studies in Language Contraction and Death*. Cambridge: Cambridge University Press.
Dorian, Nancy
1993 Internally and externally motivated change in language contact settings: Doubts about dichotomy. In: Charles Jones (ed.), *Historical Linguistics. Problems and Perspectives*, 131–154. London: Longman.
Elsig, Martin and Shana Poplack
2006 Transplanted dialects and language change: Question formation in Québec. *University of Pennsylvania Working Papers in Linguistics* 10(2): 77–90. [*Selected Papers from NWAV 34.*]
Field, Frederic
2005 Long-term effects of CS: Clues to structural borrowing. *International Journal of Bilingualism* 9(3/4): 341–360.
Gal, Susan
1979 *Language Shift: Social Determinants of Linguistic Change in Bilingual Austria*. New York: Academic Press.
Garcia, Erica C.
1985 Shifting variation. *Lingua* 67: 189–224.
Gordon, Elizabeth M., Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury and Peter Trudgill
2004 *New Zealand English: Its Origins and Evolution*. Cambridge: Cambridge University Press.
Grevisse, Maurice
1980 *Le Bon Usage*. Duculot: Paris.
Gumperz, John J. and Robert Wilson
1971 Convergence and creolization: A case from the Indo-Aryan/Dravidian border. In: Dell Hymes (ed.), *Pidginization and Creolization of Languages*, 151–167. Cambridge: Cambridge University Press.

Gutiérrez, Manuel
1994 Simplification, transfer and convergence in Chicano Spanish. *The Bilingual Review/La Revista Bilingüe* 19: 111–121.
Harris, Alice C. and Lyle Campbell
1995 *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
Haugen, Einar
1954 Review of Weinreich (1953). *Language* 30: 380–388.
Heine, Bernd and Tania Kuteva
2005 *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
Hoenigswald, Henry
1960 *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
Huffines, Marion Lois
1991 Pennsylvania German: Convergence and change as strategies of discourse. In: Seliger and Vago (eds.), 125–137.
Jakobson, Roman
1962 Sur la théorie des affinités phonologiques entre des langues. In: Stephen Rudy (ed.), *Roman Jakobson: Selected Writings*, vol. 1, 234–246. The Hague: Mouton.
Jarmasz, Lidia-Gabriela
2005 "I think I know what I'm thinking": The alternation between simple and progressive forms in stative verbs. Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
Johanson, Lars
2002 *Structural Factors in Turkic Language Contact*. London: Curzon.
Joseph, Brian D.
2007 Attitudes and ideology in contact-induced change: Trumping all psychological and structural factors. Paper presented at the 18th International Conference on Historical Linguistics (ICHL). Université du Québec à Montréal, August.
Joseph, Brian D. and Richard D. Janda (eds.)
2003 *The Handbook of Historical Linguistics*. Oxford: Blackwell.
Kaufman, Dorit and Mark Aronoff
1991 Morphological disintegration and reconstruction. In: Seliger and Vago (eds.), 175–188.
King, Ruth
2000 *The Lexical Basis of Grammatical Borrowing*. Philadelphia/Amsterdam: John Benjamins.
King, Ruth
2005 Crossing grammatical borders: Tracing the path of contact-induced linguistic change. In: Markku Filppula, Juhani Klemola, Marjatta Palander and Esa Pentillä (eds.). *Dialects Across Borders: Selected Papers from the 11th International Conference on Methods in Dialectology (Methods XI), Joensuu, August 2002*, 233–251. Amsterdam/Philadelphia: John Benjamins.
Klein, Flora
1980 A quantitative study of syntactic and pragmatic indications of change in the Spanish of bilinguals in the U.S. In: William Labov (ed.), *Locating Language in Time and Space*, 69–82. New York: Academic Press.
Labov, William
[1966] 1982 *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.
Labov, William
1969 Contraction, deletion and inherent variability of the English copula. *Language* 45(4): 715–762.
Labov, William
1972a *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.

Labov, William
1972b *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
Labov, William
1984 Field methods of the project on language change and variation. In: John Baugh and Joel Schezer (eds.), *Language in Use*, 28–53. Englewood Cliffs, N.J: Prentice Hall.
Labov, William
1994 *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.
Labov, William, Paul Cohen, Clarence Robins and John Lewis
1968 A study of the non-standard English of Negro and Puerto Rican speakers in New York City. Final report, Co-operative Research Report 3288, vol. I, US Regional Survey.
Lantolf, James
1978 The variable constraints on mood in Puerto Rican Spanish. In: Margarita Suñer (ed.), *Contemporary Studies in Romance Linguistics*, 193–217. Washington, DC: Georgetown University Press.
Lass, Roger 1980 *On Explaining Language Change*. Cambridge: Cambridge University Press.
Laurier, Michel
1989 Le subjonctif dans le parler franco-ontarien: Un mode en voie de disparition? In: Raymon Mougeon and Édouard Beniak (eds.), *Le Français Canadien Parlé Hors Québec: Aperçu Sociolinguistique*, 105–126. Quebec: Presses de l'Université Laval.
Lehmann, Winfred P.
1992 *Historical Linguistics*. 3rd ed. London/New York: Routledge.
Leroux, Martine
2003 Le passé recomposé: Analyse diachronique et multivariée de la valeur temporelle du passé en français oral de Hull (Québec). Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
Leroux, Martine
2007 Something old, something new, something borrowed, something true: What are null subjects in French? Paper presented at New Ways of Analyzing Variation 36, University of Pennsylvania, Philadelphia, October.
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* 12(2): 1–14.
Martineau, France and Raymond Mougeon
2003 A sociolinguistic study of the origins of *ne* deletion in European and Quebec French. *Language* 79(1): 118–152.
Martinet, André
1960 *Éléments de Linguistique Générale*. Paris: Colin.
Meillet, Antoine
1967 *The Comparative Method in Historical Linguistics*. Paris: Librairie Honoré Champion.
Milroy, James
1992 *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford: Blackwell.
Milroy, James
1993 On the social origins of language change. In: Charles Jones (ed.), *Historical Linguistics: Problems and Perspectives*, 215–236. London: Longman.
Montrul, Silvana
2002 Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition* 5: 39–68.
Mougeon, Raymond and Édouard Beniak
1991 *Linguistic Consequences of Language Contact and Restriction: The case of French in Ontario*. Oxford: Oxford University Press.

Mougeon, Raymond, Terry Nadasdi and Katherine Rehner
2005 Contact-induced linguistic innovations on the continuum of language use: The case of French in Ontario. *Bilingualism: Language and Cognition* 8 (2): 99–115.
Müller, Max
1871 *Lectures on the Science of Language*. 1st ed. New York: Scribner
Muysken, Pieter
2000 *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
Ocampo, Francisco
1990 The pragmatics of word order in constructions with a verb and a subject. *Hispanic Linguistics* 4(1): 87–127.
Otheguy, Ricardo, Ana Celia Zentella and David Livert
2007 Language and dialect contact in Spanish in New York: Towards the formation of a speech community. *Language* 83: 1–33.
Petrik, Katrina
2005 Deontic modality in Quebec English: "Everything you need to know". Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
Poplack, Shana
1983 Bilingual competence: Linguistic interference or grammatical integrity? In: Lucia Elías-Olivares (ed.), *Spanish in the U.S. Setting: Beyond the Southwest*, 107–131. Arlington: National Clearinghouse for Bilingual Education.
Poplack, Shana
1989 The care and handling of a mega-corpus. In: Ralph Fasold and Deborah Schiffrin (eds.), *Language Change and Variation*, 411–451. Amsterdam: Benjamins.
Poplack, Shana
1990 Prescription, intuition et usage: le subjonctif français et la variabilité inhérente. *Langage et Société* 54: 5–33.
Poplack, Shana
1992 The inherent variability of the French subjunctive. In: Christiane Laufer and Terrell Morgan (eds.), *Theoretical Analyses in Romance Linguistics*, 235–263. Amsterdam: Benjamins.
Poplack, Shana
1993 Variation theory and language contact. In: Dennis Preston (ed.) *American Dialect Research: An Anthology Celebrating the 100th Anniversary of the American Dialect Society*, 251–286. Amsterdam: Benjamins.
Poplack, Shana
1997 The sociolinguistic dynamics of apparent convergence. In: Gregory Guy, John Baugh and Deborah Schiffrin (eds.), *Towards a Social Science of Language*, 285–309. Amsterdam: Benjamins.
Poplack, Shana
2001 Variability, frequency and productivity in the irrealis domain of French. In: Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 405–428. Amsterdam: Benjamins.
Poplack, Shana
2003 Language contact and the evolution of French. Paper presented at Linguistic Symposium on the Romance Languages (LSRL) 33 and Conference on French in North America. Indiana University, Bloomington, April.
Poplack, Shana and Nathalie Dion
to appear "Prescription versus praxis: The evolution of future temporal reference in French". *Language*.
Poplack, Shana, Nathalie Dion, Lidia-Gabriela Jarmasz, Carmen LeBlanc and Nicole Rosen
2002 *Répertoire Historique des Grammaires du Français*. Corpus and documentation. Sociolinguistics Laboratory, University of Ottawa.

Poplack, Shana, Adrienne Jones, Allison V. Lealess, Martine Leroux, Chelsea T. Smith, Yukiko Yoshizumi, Lauren Zentz and Nathalie Dion
2006 Assessing convergence in contact languages. Paper presented at New Ways of Analyzing Variation 35, The Ohio State University, Columbus, November.
Poplack, Shana and Elisabete Malvar
2007 Elucidating the transition period in linguistic change. *Probus* 19(1): 121–169.
Poplack, Shana and Marjory Meechan
1998 How languages fit together in code-mixing. Special issue of the *International Journal of Bilingualism* 2(2): *Instant Loans, Easy Conditions: the Productivity of Bilingual Borrowing*: 127–138.
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(5): 707–734.
Poplack, Shana and Sali Tagliamonte
2001 *African American English in the Diaspora*. Oxford: Blackwell Publishers.
Poplack, Shana, James Walker and Rebecca Malcolmson
2006 An English "like no other"?: Language contact and change in Quebec. *Canadian Journal of Linguistics*, 51(2/3): 185–213.
Rand, David and David Sankoff
1990 *Goldvarb 2.1: A variable rule application for the Macintosh*. Montreal, Centre de Recherches Mathématiques, University of Montreal. Version 2. http://www.crm.umontreal.ca/~sankoff/GoldVarb_Eng.html.
Ringe, Don
2003 Internal reconstruction. In: Brian D. Joseph and Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, 244–261. Oxford: Blackwell.
Roberge, Yves and Nicole Rosen
1999 Preposition stranding and *que*-deletion in varieties of North-American French. *Linguistics Atlantica* 21: 153–168.
Romaine, Suzanne
1988 *Pidgin and Creole Languages*. London: Longman.
Romaine, Suzanne
1995 *Bilingualism*. 2nd ed. Malden, MA: Blackwell.
Sankoff, David
1988a Sociolinguistics and syntactic variation. In: Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, vol. 4, *Language: The Socio-Cultural Context*, 140–161. Cambridge: Cambridge University Press.
Sankoff, David
1988b Variable rules. In: Ulrich Ammon, Norbert Dittmar and Klaus J. Mattheier (eds.), *Sociolinguistics: An International Handbook of the Science of Language and Society*, vol. 1, 984–997. (Handbücher zur Sprach- und Komunikationswissenschaft 3.1.) Berlin: de Gruyter.
Sankoff, David, Sali Tagliamonte and Eric Smith
2005 *Goldvarb X: A Variable Rule Application for Macintosh and Windows*. Toronto: Department of Linguistics, University of Toronto.
Savić, Jelena M.
1995 Structural convergence and language change: Evidence from Serbian/English code-switching. *Language in Society* 24(4): 475–492.
Schmidt, Annette
1991 Language attrition in Boumaa Fijian and Dyirbal. In: Seliger and Vago (eds.), 112–124.
Schuchardt, Hugo
1884 *Slawo-deutsches und Slawo-italienisches*. Graz: Leuschner und Lubensky.

Seliger, Herbert and Robert M. Vago (eds.)
1991 *First Language Attrition.* Cambridge: Cambridge University Press.
Shapiro, Michael
1995 A case of distant assimilation: /str/ → /ʃtr/. *American Speech* 70: 101–107.
Silva-Corvalán, Carmen
1991 Spanish language attrition in a contact situation with English. In: Seliger and Vago (eds.), 151–171.
Silva-Corvalán, Carmen
1994a *Language Contact and Change.* Oxford: Clarendon Press.
Silva-Corvalán, Carmen
1994b The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change* 6(3): 255–272.
St-Amand, Anne
2002 Le subjonctif suivant une expression non-verbale. Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
St-Amand, Anne and Shana Poplack
2002 A real-time window on 19th century vernacular French: The *Récits du français québécois d'autrefois.* Paper presented at the annual meeting of the Canadian Linguistics Association, University of Toronto, Toronto.
Tagliamonte, Sali
2002 Comparative sociolinguistics. In: Jack Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), *Handbook of Language Variation and Change*, 729–763. Malden/Oxford: Blackwell.
Thomason, Sarah G.
2001 *Language Contact: An Introduction.* Washington, DC: Georgetown University Press.
Thomason, Sarah G.
2007 Language contact and deliberate change. *Journal of Language Contact* 1: 41–62.
Thomason, Sarah G. and Terrence Kaufman
1988 *Language Contact, Creolization, and Genetic Linguistics.* Berkeley/Los Angeles: University of California Press.
Torres-Cacoullos, Rena and Catherine E. Travis
to appear Variable *yo* expression in New Mexico: English influence? In: Susana V. Rivera-Mills and Daniel J. Villa (eds.), *Spanish of the Southwest: A Language in Transition.* Madrid: Iberoamericana/Vervuert.
Toribio, Almeida Jacqueline
2004 Convergence as an optimization strategy in bilingual speech: Evidence from codeswitching. *Bilingualism: Language and Cognition*, 7(2): 165–173.
Van Coetsem, Frans
2000 *A General and Unified Theory of the Transmission Process in Language Contact.* Heidelberg: Winter.
Weinreich, Uriel, William Labov and Marvin Herzog
1968 Empirical foundations for a theory of language change. In: Winfred P. Lehmann and Yakov Malkiel (eds.), *Directions for Historical Linguistics*, 97–195. Austin: University of Texas Press.
Whitney, William Dwight
1881 On mixture in language. *Transactions of the American Philosophical Association* 12: 1–26.
Willis, Lauren
2000 *Être ou ne plus être*: Auxiliary alternation in Ottawa-Hull French. Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
Winford, Donald
2003 *An Introduction to Contact Linguistics.* Oxford: Blackwell.

Winford, Donald
2007 Some issues in the study of language contact. *Journal of Language Contact – Thema 1 (2007)*. Available online from <http://www.jlc-journal.org>.
Zentz, Lauren
2006 "C'est ça je travaille dessus": Orphaned prepositions and relativization in Canadian French. Unpublished MA thesis (*mémoire*), Department of Linguistics, University of Ottawa.
Zribi-Hertz, Anne
1984 *Orphan Prepositions in French and the Concept of "Null Pronoun"*. Indiana University Linguistics Club: Bloomington, IN.
Zwicky, A.
2005 More illusions. Available at <http://itre.cis.upenn.edu/~myl/languagelog/archives/002407.html>. Accessed 29 April 2008.

Shana Poplack and Stephen Levey, Ottawa (Canada)

23. Areal language typology

1. Introduction: Setting the stage

Linguistic typology deals with linguistic variation and the discovery of cross-linguistic universal patterns reflected in that variation. The aim of the present article is to show that the robustness and the reliability of typological generalizations crucially depends on the appropriate assessment of areality, i.e., the influence of geographic closeness and language contact on structural properties of languages. Thus, the analysis of structures and their overall statistical distribution in the population of the world's languages does not guarantee any direct access to valid generalizations about language universals and about properties of the human cognition.

This can be illustrated by an example from basic word order. In his 1986 volume, Tomlin looked at the overall percentage of the six logically possible combinations of S (Subject), O (Object) and V (Verb) in a statistically balanced sample of 402 languages. As it turned out, the majority of these languages were either SOV (180 languages, 44.78 percent) or SVO (168 languages, 41.79 percent). Only 37 languages (9.20 percent) were VSO. The sequences in which O preceded S were only represented by a small minority of twelve languages (2.99 percent) with VOS, five languages in the Amazonian region of