Language Variation and Change, **21** (2009), 297–317.
© Cambridge University Press, 2009 0954-3945/09 $16.00
doi:10.1017/S0954394509990196

Replication, transfer, and calquing: Using variation as a tool in the study of language contact

MIRIAM MEYERHOFF
University of Edinburgh

ABSTRACT

Do the processes of replication, transfer, and calquing operate on speakers' mental organization of variables? Can the comparison of constraint rankings across languages provide evidence for (or against) the transfer of features in cases of long-term language contact? This article suggests yes to both questions. It undertakes a comparison of constraint rankings on null subjects and null objects in a corpus of Bislama and a corpus of Tamambo narratives to demonstrate the potential and limitations of such methods. It concludes that these methods: (i) allow us to particularize our definitions of *replication*, *transfer*, and *calquing* to inherently variable domains, and (ii) strengthen connections between variationist sociolinguistics and the fields of creolistics and language contact.

This article has two goals. First, it explores whether the processes of transfer and calquing (commonly discussed in studies of language contact, and specifically creolization) can be discerned in patterns of variation. That is, does the process of calquing operate outside the canonical domains of morphology and syntax and in relation to speakers' mental organization of variables?

Second, it explores whether the increasingly common use of comparisons of constraint rankings for the same (or ostensibly the same) variables in different language varieties can be used as inferential evidence for (or against) the transfer of features in cases of long-term language contact.

The distinction between weak transfer, strong transfer, and calquing was first presented at conferences in York, Paris, and Philadelphia in September/October 2007. I thank audiences at the Association of Linguistic Typology, United Kingdom Language Variation and Change, and New Ways of Analyzing Variation conferences, as well as post-graduate students in my Sociolinguistics Reading Group 2008–2009 for their comments and questions, in particular, Isabelle Leglise, Isabelle Buchstaller, Karen Corrigan, Dave Embick, and Scott Schwenter. Dave Britain and Barbora Skarabela gave comments on an earlier draft. Gregory Guy, Naomi Nagy, John Paolillo, Robert Sigley, Gerard van Herk, and James Walker all chipped in with advice on conducting cross-linguistic comparison of variation, some of which has not been possible to adopt, but all of which I am grateful for (*ol tawe blong mi ia*). I am especially grateful to James Stanford and Dennis Preston who urged me to do a detailed comparison of Tamambo and Bislama in the first place. The Wenner-Gren Foundation funded fieldwork in 1994–1995, during which the Bislama data was collected. Tamambo data is courtesy of the ever-collegial Dorothy Jauncey. This article has benefited enormously from constructive and detailed criticism by two LVC reviewers: thank you both. None of the above are responsible for remaining errors and infelicities.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

This article will propose that the two goals can be linked methodologically, and that the analysis of variation:

a. can enrich our understanding of what happens in situations of language contact, and
b. enables us to particularize how we define replication, transfer, and calquing to inherently variable domains.

I begin with some general background comments on the two goals of the article, and then outline the general problem for creolistics in particular. I then present a variationist framework for systematizing the investigation of (a) and (b). I then review how such a framework deals with some data that relates to the problem I have outlined. Finally, I suggest some implications of this approach for the fields of variationist sociolinguistics and creolistics.

SOME COMMENTS ON REPLICATION, TRANSFER, AND CALQUING

Research on language contact recognizes that when speakers of different language varieties are in more than superficial contact with each other, there is the potential for transfer on all levels of linguistic structure. Some researchers propose that there are hierarchies governing how easily transfer takes place on different structural levels—hierarchies that make more or less reference to social and psychological factors as well as language-internal ones (Heine & Kuteva, 2005; Matras, 2000; Thomason & Kaufman, 1988; Winford, 2003, all review this approach in the development of the field of language contact).

Among the effects of transfer over time in communities and within the repertoire of individuals, the *calquing* of linguistic features from one language to another is well-documented. A *calque* can be defined quite specifically in linguistic terms, and in the field of language contact, a *calque* is usually used to refer to the direct translation, morpheme by morpheme, or word for word, of concepts and syntactic structures that originated in one language and can be shown to be (or be argued to have been) a historical introduction into another. The realization of the morphemes in English *skyscraper* (sky + scraper) in German as *Wolkenkratzer* (*Wolken* 'cloud' + *Kratzer* 'scratcher') is often cited as an example (e.g., Winford, 2003:44–45). Semantic equivalence also underlies the use of the term calque to describe cross-linguistic influences from one language to another on syntax and multiword constructions (e.g., Hock, 1992; Turell, 1998).

Heine & Kuteva (2005:ch. 2, especially) provided a host of examples of the more general process that they prefer to call *replication*, where there is a *model* language in which a pattern originates historically, and a *replica* language, which adopts the pattern as a result of contact with the model. This terminology is helpful, and I will follow them in their use of model and replica languages to

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

identify which of the varieties in contact we are talking about with respect to a particular pattern. Describing the underlying process as replication (rather than as transfer or interference) has some advantages too, because replication suggests both copying and differentiation. Transfer, by contrast, suggests relocation of a feature or subsystem in toto, and interference suggests problems with respect to the target. In most of the cases Heine & Kuteva discussed, replication does involve some nativization or alteration of the model in the replica. The data I will introduce later illustrate this nicely.

Heine & Kuteva discussed examples of replication that include the development of a "whether" complementizer into a polar question marker in Finnish Swedish, apparently on the model of a Finnish pattern (Koptjevskaya-Tamm & Wälchli, 2001); the use of "do"-support in varieties of Pennsylvania German (Burridge, 1995); and the conventions for use of cleft constructions in Irish English (which replicate the pragmatic constraints on clefting in Irish more closely than the pragmatic constraints on clefting in English [Cotter, 1994]). Similar morphosyntactic examples of transfer (e.g., what has been referred to as a calquing of *back* from English into Ontario French) are discussed by King (2008) for French-English contact. In the field of creolistics, many claims have been made that structures in the emergent creoles are the result of calquing or morphosyntactic transfer (Camden, 1979, and Keesing, 1988, being particularly radical proposals of this for Bislama and Solomon Islands Pijin, respectively). We will consider the case of language contact and creolization in more detail shortly.

In all these examples, it is worth noting the obvious. The process involves the transfer of semantics, with either the attendant transfer of form (phonetic realization) from one language to another, or the substitution or reallocation (cf. Britain & Trudgill, 2005; Trudgill, 1986) of a semantically equivalent (or related) form in the replicating language.

But notice that there is some indeterminacy in application of the definition in some case studies. For example, with *Wolkenkratzer*, the English morphemes are not directly translated (**Himmelschaber*); instead suitable candidates within the wider semantic field of {sky} and {scrape} are selected.[1]

Even looser semantic mappings may emerge at the level of structure. This is perhaps particularly true when we consider the very profound language contact involved in the formation and development of creoles. For example, a distinction between first-person plural (1p) inclusive and 1p exclusive pronouns is found in all varieties of Melanesian English. This is said to be the result of the transfer or calquing of a semantic distinction that is found in the substrate languages. What is interesting about this case, is that the purported calque is realized in the replica with morphemes taken from the model language, even though the model has no direct analogue of the inclusive/exclusive distinction. This can be seen in Table 1, which compares the pronouns in Panen (a very typical Vanuatu language, spoken in Matantas in Big Bay [Meyerhoff, 2008]) with the pronouns in Bislama (the creole, national language of Vanuatu). Consider particularly, the first- and second-person forms.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

TABLE 1. *Independent pronouns in Panen (spoken in Matantas, Vanuatu) and in Bislama (national language creole)*

Independent pronouns (Panen)				
	1 inclusive	1 exclusive	2	3
Singular	—	inau	niho	inie
Plural	hitse	kanam	kaniu	ihire
Independent pronouns (Bislama)				
	1 inclusive	1 exclusive	2	3
Singular	—	mi	yu	hem
Plural	yumi	mifala	yufala	olgeta

It is clear that whereas the Bislama 1p pronouns are (a) derived, and (b) semantically transparent, the same cannot be said of the Panen 1p pronouns (any historical compositionality in the Panen pronouns is synchronically opaque). In other words, the sense of the distinction (inclusive versus exclusive) may have transferred, but the mechanics for drawing that distinction are not what was transferred.

This sort of creative mapping in creoles (and its implications for genetic theories of creole evolution) is discussed at length in Mufwene (2001). Such loose mappings in the process of calquing may be particularly typical of the Pacific creoles, because, as we know now, they were all characterized by extensive communal and individual bi-/multilingualism in their history (this is uncontroversial for the Melanesian creoles—and continues to be the case for many speakers; Roberts' [2000] work on Hawai'i Pidgin convincingly makes the same case for extended periods of bilingualism there).

Clearly, it is uncontroversial to assert that there is a lot of variability in the mechanisms and hence the final outcomes of language contact. But what kind of place does inherent variability—the inter- and intraspeaker variation characteristic of all speech communities—have in the field of language contact? In the next section, I offer some remarks about this and move toward a proposal for a systematic framework under which variation can be linked to wider trends in the study of language contact.

SOME COMMENTS ON VARIATION, LANGUAGE CHANGE, AND LANGUAGE CONTACT

By and large, work on language contact acknowledges the relevance of differences in frequency and of variation (in either the input or the output). Heine & Kuteva discussed the shift from what they call "minor use patterns" to "major use patterns" (2005:44ff) and frequency changes with replication. Thomason & Kaufman (1988), Trudgill (2004), and Siegel (1999) all note that the interplay of

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

different variants in the input may be crucial to the output in situations of radical intervariety contact and creolization. However, in general, work on creole (or mixed language) genesis acknowledges that differences in frequency and/or variable input and output are relevant, without making them the explicit focus of enquiry.

There are notable exceptions to this. Recent years have seen the increased use of variation itself as the focus of comparative historical work and in the exploration of cross-linguistic transfer. Pioneered by Shana Poplack and her associates, the comparison of significant linguistic constraints on the variation between the same forms in different varieties of English, or Romance languages (e.g., Poplack, 2000; Poplack & Malvar, 2006; Poplack & Tagliamonte, 2001; Tagliamonte, 2002) has been used to infer historical relatedness (where there are similar significant factors) or, conversely, a lack of historical relatedness of language varieties (where the significant linguistic factors constraining variation are different). Using these methods, Naro & Scherre (2000) argued that some variables in Brazilian Portuguese reflect inherited traits from peninsular Portuguese.

Some quantitative studies of language contact have also explored variation as diagnostics of transfer and replication. G. Sankoff (1993) examined the gradual change in preferred patterns of use of *yet* in Tok Pisin. She found that over time the pattern looks less and less like the English distribution of *yet* and more and more like the distribution of Tolai *iat*, eventually resulting in a synthetic pattern that (she argued) reflects the extended temporal contact between English and Tolai (and perhaps other vernacular languages). A more detailed use of the comparison of variation in cases of language contact is Blondeau & Nagy's (2008) work on subordinate clause marking in L1 and L2 Canadian French and English. Buchstaller (2008) and Buchstaller & D'Arcy (2009) used the same methods for dialect contact. Meyerhoff & Walker (2007, 2009) used the comparison of constraints on variation to argue that certain individuals in a community have not changed their grammars over time, despite superficial differences in the frequency with which they use "local" variants.

G. Sankoff's work remains (to my knowledge) the only example of the use of variation to address questions of the emergence of structure in creoles (but see Britain, 2008; Kerswill & Williams, 2005; and Trudgill, 2004, on new dialect formation). However, this article will suggest that unless we seriously evaluate variation (which so many researchers agree is an important component in the overall picture of creole development), we may erroneously attribute inheritance, or mistakenly infer historically shared development, to varieties that are not in fact "sisters under the skin."

Variationists are well-equipped to make a helpful contribution to this. In the next section, I briefly outline the case for focusing on variation in the genesis and development of creoles and other contact languages. I then offer a proposal for using quantitative methods—familiar to variationists—in the service of providing an objective and data-driven basis for discriminating between different kinds of contact-induced effects. Ultimately, this provides a basis for discriminating among replication, transfer, and a highly restrictive notion of calquing.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

A CREOLIST'S PROBLEM: DOES VARIATION REPLICATE?

It is reasonably well-attested that, at least for the Melanesian creoles, some categorical features of the substrate languages appear as variable or stochastic features of the creole (Meyerhoff, 2000, 2003, for Bislama; G. Sankoff & Brown, 1976; G. Sankoff & Mazzie, 1991; Smith, 2002, for Tok Pisin). In fact, if the competition between lexifier models and substrate models for any given construction is seen as a form of variation, then we can say it has long been recognized that variation in the input is a crucial aspect of how creoles develop. G. Sankoff (1994), Crowley (1990), and Siegel (1999) all argue that a synthetic process that resolves substrate, lexifier, and cognitive factors is central to the development of the Melanesian creoles (and Mufwene, 2001, for creoles more generally). Meyerhoff (2000:224) suggested that "convergence between the lexifier and the substrate is the proximate cause of creole variation." Mufwene (1996) proposed that the specific role of Universal Grammar in such linguistically dynamic situations is to predict which variants, out of the mix, will be focused (Le Page & Tabouret-Keller, 1985) in the resulting creole.

Other work has proposed even more direct links, suggesting that variable features of the substrate(s) may be the source of the variation observed in creole grammars. Meyerhoff (2000) noted that the typologically unusual pattern of pro-drop in Bislama (*pro* favored in third person, and pronouns favored in first and second) may have syntactic models in the substrate languages of Vanuatu, but that the necessary data on this kind of variation are usually unavailable in reference grammars.

Jauncey (1997) noted that in the Vanuatu language Tamambo some nouns can be treated as alienable or inalienable in different contexts, that is, alienability is a property that emerges in discourse, and she intimates that there may be inherent variability in possession marking, but that the contextual constraints on this are difficult to pin down. Following from Jauncey's observation, Meyerhoff (2003) suggested that the probabilistic marking of alienability in Bislama object NPs may be linked to its variability in (some) substrate languages. Crowley (1990, 2004) argued that animacy, a probabilistically marked feature of substrate grammars, is the source of variation in the form of Bislama subject and object NPs (null versus overt) and that it also underlies the alternation between the agreement marker *i* and *oli* with third-person plural subjects (*oli* is canonical; *i* can be analyzed as the default agreement marker in Bislama [Meyerhoff, 2001]).

There is no a priori reason why replication of variable features from the substrate into a creole might not occur, and as the preceding examples demonstrate, there have been consistent (if muted) claims to this effect among researchers studying creoles for some time. From a broader typological perspective, there is even less reason to suppose this might not be the case. Other work on language contact has demonstrated solid evidence for the transfer of variable features of one language or variety into another (e.g., Blondeau & Nagy, 2008).

Despite this, until now, there has been limited direct quantitative comparison of variation in substrate languages and their related creoles. Drawing on some relevant

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

data from Bislama (the creole spoken in Vanuatu, Southwest Pacific), this article proposes a way of typologizing different levels of replication or transfer within a variationist perspective.

A VARIATIONIST'S APPROACH TO SOLVING THIS PROBLEM

I draw on the comparative variationist approach established by Shana Poplack. This involves subjecting data from different groups of speakers to the same multivariate analysis. The factor groups and factors within groups are held constant across the different populations, and the resulting Goldvarb (Goldvarb X [D. Sankoff, Tagliamonte, & Smith, 2005]) weightings can be compared. Differences in which factor groups are returned as significant, and differences in the ranking of factors within those groups, are interpreted as diagnostics of fundamentally different underlying grammars of variation. (Poplack's approach advances a clear position with respect to the status of variation in the grammar, viz. variation is emphatically a property of human grammar, rather than being an artifact of pressures peripheral to the core grammar.)

In Meyerhoff & Walker (2007), it became clear this method is well-suited to capture the gradient effects associated with contact-induced transfer. In particular, the basic components of a variationist analysis—identification of significant factor groups, identification of the most parsimonious analysis of factors within groups, and the different weighting of factors within groups—lend themselves to three distinctions. These are outlined in (1).[2]

(1) *Principles for incorporating variation into the study of language contact:*

(i) Where the same factor groups are significant constraints on a variable in the model and in the replica varieties, let us call this *weak transfer* or *replication*;
(ii) Where the same factor groups are significant in both model and replica, *and* the ordering of these factor groups is the same in both model and replica, let us call this *(strong) transfer*;
(iii) Where the same factor groups are significant in both model and replica, *and* the ordering of these factor groups is the same in both model and replica, *and* the factors *within* groups have the same ranking in model and replica, let us call this *calquing*.

This typology considers similarities both of kind and of quality. In doing so, it allows us to ask linguistically and statistically meaningful questions about whether the same surface-level form found in two languages is in fact "the same." It adopts well-established terms from the field of language contact and attempts to relativize them to the kind of phenomena that variationists study.

Table 2 summarizes these principles schematically, in the manner of Buchstaller & D'Arcy (2009). Because Buchstaller & D'Arcy are concerned with dialect contact, their focus is somewhat different. They include a category for replication of form only (with no associated sociolinguistic constraints)—

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

TABLE 2. *Operationalizing replication, transfer, and calquing with variables*

	Same factor groups in both Ls	Same order of factor groups in both Ls	Same order of groups and factors *within* groups in both Ls
Weak transfer/ replication			
Strong transfer			
Calquing			

dialects of the same language are not usually dealing with the problem of replicating an underlying semantics, so the option of form-only replication in fact amounts to form and (I assume, truth-conditional) semantics.

In some respects, what we are trying to do is systematize the approximate quality of transfer and mapping between varieties that was exemplified earlier with the example of English *skyscraper* being calqued into German as *Wolkenkratzer*. However, it is important to note that the purpose of this article is not to provide a basis for classifying pairs such as *skyscraper-Wolkenkratzer* as calques or not. The typology proposed in (1) is clearly framed so that it is concerned with cases in which inherent variability is a feature of both the model and the replica language. Pairs such as *skyscraper-Wolkenkratzer* do not involve inherent variability.

In the next section, I demonstrate how the framework in (1) might be used in the study of a particular case of language contact, drawing on data from an indigenous language of Vanuatu, Tamambo, and Vanuatu's national language, the creole Bislama.

DATA SET AND METHODS

In this section, I revisit two of the Bislama variables already mentioned—phonetically null subjects and phonetically null objects—and I compare them directly to analogous variables in Tamambo, the language spoken on Malo island in northwest Vanuatu.

The motivation for selecting Tamambo is both principled and pragmatic. First, many of the speakers in my corpus of conversational Bislama were L1 speakers of Tamambo or of closely related languages from South Santo or Ambae. Hence, the corpus is a good candidate for testing the possible transfer of constraints into the Bislama of the corpus that is also available. Second, Dorothy Jauncey, who has done extensive work on Tamambo, kindly offered me access to a corpus of Tamambo texts she had collected in 1994–1995. This was coincidentally the same time I was collecting my Bislama corpus, however, none of the speakers in her Tamambo texts overlap with speakers in my Bislama

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

corpus. This is important because it means any parallels we might identify cannot be attributed to the transfer of idiosyncratic norms or speaker preferences.

Jauncey (2002) provided a structural overview of Tamambo.[3] It is a very typical, rather conservative member of the North-Central Vanuatu family. From her corpus of texts, I selected the *tandono* (traditional, customary stories) for analysis. As narratives, these were closer in genre type to the narratives and free conversation of the Bislama corpus. The narratives were told aloud and subsequently transcribed with the help of (different) native speakers. The transcription seems to try and reproduce the spoken quality of the narratives, for example, often marking apocope. It is much smaller than the corpus of Bislama spoken by Malo speakers, but as we will see, it nonetheless allows us to pose and answer the question of whether variable properties of the substrate show up in the creole.

There are several constraints that are potential transfer candidates and which we would therefore like to investigate in both corpora. As noted, Crowley (2004) proposed that animacy of the referent was a constraint on the variable occurrence of subjects and objects in the substrate languages of Vanuatu. He proposed that this was the source of the subject and object variation observed in Bislama. However, his quantitative work on this variable considered only third-person subjects (Crowley, 2004; see Meyerhoff, 2009, for details), a definition of the envelope of variation that is perhaps too narrow, as it will necessarily fail to reveal a continuum of internal differences between the way animate third-person referents are treated and the way inherently animate first- and second-person referents behave. Consequently, we will want to consider animacy in a broader sense—across all persons—in the Tamambo and Bislama corpora.

Other work on the distribution of overt and phonetically null subjects and objects in Bislama has identified two other factor groups that are likely to be relevant and helpful in evaluating possible transfer of variable components of the grammar. These are the form and grammatical role of the argument's referent in the immediate discourse (for both subjects and objects), and whether or not the referent would be considered a prototypical alienable or inalienable possession (e.g., in a possessive NP). Discourse factors are relevant for both subjects and objects (Meyerhoff, 2000, 2003), and (in)alienability has been found to be relevant for objects (Meyerhoff, 2003). Both (in)alienability and form and grammatical role in the prior discourse will be compared in the Bislama and Tamambo corpora.

Bislama and Tamambo share many typological features, but their morphology and syntax diverge in some key ways. This requires us to take a catholic approach to the analysis of objects. In both Tamambo and Bislama, subject arguments allow the same alternations: phonetically null versus overt (pronoun or NP) realization of the subject. Both languages also have some overt subject agreement morphology associated with the verb (whether these morphemes are analyzed as agreement or as bound/clitic subjects is immaterial to this investigation), and this morphology is a potential source of identification of the referent when there is no phonetically overt subject. But in object position,

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

Tamambo does not have the option of a zero variant as Bislama does. There is a phonetically reduced variant, but this is an object clitic that appears on the verb (it should not be compared with the Bislama transitive suffix, *-Vm*, which does not seem to identify any particular argument, e.g., third person [Meyerhoff, 1996]).

Given that the independent pronoun–NP/clitic alternation in Tamambo fills the same kind of ecological niche that the pronoun/zero alternation fills in Bislama, it is reasonable to compare the distribution of these variants in the two languages (such concessions are required frequently when undertaking cross-linguistic comparisons of phenomena at the syntax-discourse interface. Gundel, Hedberg, & Zacharski [1993] use similar equivalences as the basis for their comparison of the realization of given/new information in a variety of languages).

As noted, the corpora are not matched for size. The analysis of subjects in Bislama is based on the speech of 18 speakers from Malo island between the ages of 10 and early 40s. This corpus contains 1861 clauses with possible subjects (all nonfinite verb phrases were excluded, likewise finite verb phrases where a subject is not possible, i.e., existential, weather verbs, second verbs in serial verb constructions, see Meyerhoff [2000] for details). The corpus of Bislama objects is necessarily smaller, because not all verbs are transitive. The analysis of objects in Bislama is based on the speech of 13 speakers from Malo island and contains 869 clauses with possible objects.[4]

The Tamambo corpus is *tandono* stories told to Dorothy Jauncey by two elders in West Malo, Vula Vutilolo and Vira Joseph, transcribed and translated into English with the assistance of Selina Vutilolo and Lucian Bani. The texts are not given morpheme by morpheme glosses; I have used my own knowledge of Tamambo and Jauncey's (1997) grammar. The corpus consists of 288 clauses, with 274 possible subjects and 147 possible objects.

The dependent and independent variables coded are summarized here.

Dependent variables:

- Bislama subjects: NP (2) and (3), free pronoun (2), no free pronoun or NP (3):

(2)

Sapos	wan	man	i	save	faenem	wan	wok	blong	mi
If	a	man	AGR	can	find	a	work	of	1s
O	mi	aot	long	ples	ya				
Oh	1s	out	from	place	SPEC				

'If someone found me a job, oh, I'd be out of here.' (Dien, M-95-12)[5]

(3)

ol	boe	oli	krangke	olbaot	nomo
PL	boy	AGR	crazy	all.about	only
Ø	oli	stap	dring		
Ø	AGR	CONT	drink		

'The boys were just completely crazy, (they) were drinking.' (Jackson, M-95-15)

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

- Tamambo subjects: NP (5), free (independent) pronoun (4), no free pronoun or NP (4) and (5):

(4)

Moiso	kumboi	ku	storihi	na	tamalohi	atea.
so	1s.FUT.like	1s	story	a	man	one
Nia	le	losu	tambaluhina	asena.		
3s.INDEP	ASP	beat	wife.3s.POSS	very		

'So I would like to tell a story about a man. He used to beat his wife a lot.' (*Manja*, Vira Joseph, 5-9-94)

(5)

Uita	nia	le	ate.	Mo	duhu	a	le	ate
octopus	3s.INDEP	ASP	be.there	3s	good	3s.IRR	ASP	be.there

'The octopus is there. It's alright there.' (*Tarikoko*, Vula Vutilolo, 6-9-94)

- Bislama objects: NP (6), free pronoun, no free pronoun or NP (6):

(6)

Afta	Ø	oli	karemaot	Rambo	long	prisin	
then	Ø	AGR	take.out	R	from	prison	
Ø	i	kam	kilim	Ø,	Ø	i	ded
Ø	AGR	come	kill	Ø	Ø	AGR	dead

'So (they) get Rambo out of prison, (and) (he) comes kill (him), (and then) (he) is dead.' (Bretian, M-95-10)

- Tamambo objects: NP or free pronoun, phonetically reduced/clitic object (NP and clitic shown in (7)):

(7)

Bao	mo	lako	Tao	talom.	Mo	lakoa	mo	duhu
swamphen	3s	decorate	rooster	first	3s	decorate.3s	3s	good
asena.								
very								

'Swamphen decorated Rooster first. He decorated him really well.' (*Bilahe, Toa, Bao*, Vira Joseph, 5-9-94)

Independent variables:

- Person and number of referent: first, second, third person; singular, plural
- Animacy: speaker and addressee, third-person human, third-person animate (nonhuman), third-person inanimate
- Grammatical role and form of the referent in immediately prior discourse: overt subject, null subject, overt other argument, null other argument, not present in prior clause.

Previous work on the Bislama corpus has indicated that perseverance of the referent across clauses operates very locally, so for this reason the coding refers strictly to the prior clause. Note that, defined like this, "not present in prior clause" includes subject reference-switching between clauses, as well as brand new referents in the discourse.

It is possible to make finer-grained distinctions for both animacy and grammatical role/form of referent in the Bislama corpus, but because the Tamambo corpus is

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

smaller, this represents the optimal set of factors and factor groups for comparison. Person and number of the referent and animacy are not wholly independent variables, but we will see that it appears to be possible to identify independent effects.

For each dependent variable, the reduced variants (null subjects in Bislama and Tamambo; null objects in Bislama and clitic objects in Tamambo) were treated as the application value for a multivariate analysis using Goldvarb X (D. Sankoff et al., 2005).

I turn now to a report of the results from that investigation. I then proceed to an evaluation of the findings in the light of the typology in (1).

THE SUBJECT VARIABLE IN TAMAMBO AND BISLAMA

Table 3 compares the constraints on the form of the subject in both the Tamambo and Bislama corpora. In both languages, the principal constraint is one that essentially reflects information status of the referent: grammatical role and form of the referent in the prior clause. In both languages, argument referents that were not overt in the preceding clause favor a nonovert form as the subject of the current clause. In Tamambo, these are the only favoring factors. In Bislama, there is also a preference for continuing subjects (whether overt or nonovert in the prior clause) to be null in the current clause. At the other end of the scale, if the referent was not mentioned in the last clause, then this strongly favors some form of overt realization for the current subject. A constraint of discourse salience operates universally in some form or another on null or reduced/clitic subjects.

In Tamambo, the only other significant constraint on the form of the subject is animacy. Person and number of the subject is not significant (though some cells have very small numbers). In the majority of clauses, the agreement morphology that occurs with a finite verb in Tamambo uniquely identifies the subject in virtually all persons and numbers (this is true in realis clauses and nonfuture irrealis). The only allomorphy is between 1p inclusive and 1p exclusive, so this finding is a little unexpected (the full verb paradigm can be seen in the appendix). A common, long-held account for the distribution of null subjects is that they are favored cross-linguistically where the verb morphology provides reliable cues to the subject reference (Jaegli & Safir, 1989).

Conversely, person and number of the subject referent is a significant constraint in Bislama and animacy is not significant. The ranking of different persons and numbers departs slightly from the pattern found in the larger Bislama corpus discussed in Meyerhoff (2000); however, when we look at percent null subjects, we see the same pattern discussed there, a clear preference for overt subjects in first and second person, and higher frequencies of nonovert subjects in third person.

It is interesting to consider the (nonsignificant) results for animacy in Bislama in a little more detail. The difference in factor weights for this group is quite large for the Bislama data, and as noted earlier, animacy and person/number of the subject are not wholly independent. It is possible that the reason animacy fails to emerge as a significant factor group for Bislama is related to the overlap between the highest

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

TABLE 3. *Comparison of constraints on absence of free N/pronoun subjects in Tamambo and Bislama*

	Tamambo			Bislama		
	Total = 272 Input = 0.81 Factor weight	%	*n*	Total = 1837 Input = 0.29 Factor weight	%	*n*
Grammatical role and form of referent in prior clause						
Subject (not overt pronoun/N)	0.74	92	148	0.91	85	116
Other argument (not overt pronoun/N)	0.69	88	16	0.73	47	19
Subject (overt)	0.34	63	19	0.64	30	786
Other argument (overt)	0.26	53	15	0.50	32	133
Not present	0.13	41	74	0.29	18	781
Range	61			62		
Animacy						
1st/2nd person	0.57[a]	80	26	[0.21]	[05]	823
3rd human	0.60	82	147	[0.76]	[52]	787
3rd animate (other)	0.37	65	74	[0.70]	[57]	14
3rd inanimate	0.20	44	25	[0.71]	[33]	213
Range	40					
Person and number of subject referent						
1 singular	[0.71]	[88]	24[b]	0.35	04	451
2 singular	[0.51]	[50]	2	0.45	06	215
3 singular	[0.46]	[70]	161	0.46	38	666
1 plural	n.d.[b]	n.d.[b]		0.48	06	137
2 plural	n.d.	n.d.		0.58	05	20
3 plural	[0.52]	[76]	85	0.78	69	348
Range				43		

[a]Ranking of factors within *Animacy* factor group suggests human versus nonhuman distinction in Tamambo; discourse participants versus nonparticipants distinction in Bislama. [b]Four (null) tokens of 1p referents were recoded with 1s tokens.

animacy speaker-addressee and first- and second-person forms. We can see that the animacy split in Tamambo is essentially between human and nonhuman referents (human referents favoring null variants), whereas in Bislama, the split is between first- and second-person referents and all other referents, that is, discourse participants (which favor a free pronoun or NP realization) and nonparticipants (which favor null forms).[6]

THE OBJECT VARIABLE IN TAMAMBO AND BISLAMA

Table 4 compares the constraints on the form of the object in the Tamambo and Bislama corpora. As with subjects, the principal constraint in both languages is one of discourse continuity. In Tamambo, if the object's referent had been a thematic argument in the prior clause, this favored its realization as a clitic. In

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

TABLE 4. *Comparison of constraints on absence of free N/pronoun objects in Tamambo and Bislama*

	Tamambo			Bislama		
	Total = 147 Input = 0.40 Factor weight		*n*	Total = 869 Input = 0.14 Factor weight		*n*
		%			%	
Grammatical role of referent in prior clause						
Subject	0.89	77	22	0.82	37	93
Direct object	0.87	81	37	0.90	62	217
Other argument	0.35	30	10	0.70	28	195
Not present	0.20	17	78	0.11	02	364
Range	69			79		
Alienable/inalienable possession[a]						
Inalienable	0.71[b]	54	39	0.35	21	134
Alienable	0.42	38	106	0.53[b]	27	735
Range	29			18		
Animacy						
Human referent	[0.39]	[47]	32	[0.43]	[27]	195
Other animate	[0.58]	[48]	29	[0.68]	[51]	51
Inanimate	[0.52]	[40]	86	[0.51]	[24]	623

[a]Two tokens of 1s objects not coded for alienability. [b]Reversal of effect of *(in)alienable possession* in Tamambo and Bislama.

Bislama, the distinction is less subtle. If the object referent had been mentioned at all in the prior clause (regardless of its grammatical role), this favored a phonetically null realization in the current clause.

Both languages also show a significant effect for whether or not the object refers to something that would be considered an alienable or inalienable possession. However, in the two languages, the direction of the effect differs. Inalienable direct object referents are more likely to be reduced (clitics) in Tamambo; in Bislama, alienable objects are more likely to be reduced (phonetically null).

Animacy of the referent is not a significant constraint on the presence or absence of a free pronoun/full NP in either language (and, indeed, the ranking of individual factors within this group does not look like any normal animacy hierarchy, except perhaps an incipient human/nonhuman distinction for Tamambo).[7]

Table 5 summarizes the relative strength of the significant constraints for subjects and objects in the two languages. We can now explore to what extent the distinctions proposed in (1) can be usefully applied to these data.

APPLICATION OF THE PRINCIPLES OF REPLICATION, TRANSFER, AND CALQUING

The proposed typology for incorporating variation directly into the discourses and traditions of language contact research draws on distinctions we can make about

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

TABLE 5. *Summary of significant constraints on the subject and object variables in Tamambo and Bislama*

	Significant constraints	
	Tamambo	Bislama
Subjects	1. Form/role of referent in prior clause 2. Human vs. nonhuman referent	1. Form/role of referent in prior clause 2. Person/number
Objects	1. Grammatical role of referent in prior clause 2. Inalienable possession	1. Grammatical role of referent in prior clause 2. Alienable possession

Note: Order reflects strength of the factor groups within each variable.

how similar a variable is in both the model and replica languages, where similarity can be measured in quality and quantity. For convenience, I reproduce (1) here.

(1) *Principles for incorporating variation into the study of language contact:*

(i) Where the same factor groups are significant constraints on a variable in the model and in the replica varieties, let us call this *weak transfer* or *replication*;
(ii) Where the same factor groups are significant in both model and replica, *and* the ordering of these factor groups is the same in both model and replica, let us call this *(strong) transfer*;
(iii) Where the same factor groups are significant in both model and replica, *and* the ordering of these factor groups is the same in both model and replica, *and* the factors *within* groups have the same ranking in model and replica, let us call this *calquing*.

Under this proposal, *weak transfer* or *replication* of variables (which in all likelihood is the most common) involves cases where a feature occurs variably (noncategorically) in both model and replica, and there is overlap in the inventory of linguistic factors that are significant constraints on the variable in both varieties.[8] However, what may be the strongest constraint on the variable in the model language may be the least strong constraint on the variable in the replica language.

We have seen a partial illustration of this replication, or weak transfer, in the data on subject deletion. The factor groups that significantly constrain absence of a free nominal or pronominal subject differ in Tamambo and Bislama. In Tamambo, the most significant factor is the referent's form and role in the prior clause, followed by animacy of the referent. In Bislama, the most significant factor is the referent's form and role in the prior clause, followed by person and number of the subject. If Bislama has replicated the model found in Tamambo patterns, the process only partly replicates the significant factor groups in the model.

In addition, we would not want to place too much emphasis on the congruence of the first factor group. As noted, it is cross-linguistically unremarkable for

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

null subjects to be constrained by the discourse salience of the referent. This general fact about the distribution of null subjects should make us cautious about how much of the parallelism between Tamambo and Bislama we attribute to pattern replication. The point is that if we compared any two null subject languages, whether in contact or not, we would expect to find that they share a constraint on the distribution of the null variants that is discourse-pragmatic.

Within the factor group that describes the form and role of the referent in the prior clause, we see rankings of individual factors that are (again) unsurprising cross-linguistically, however, we did note an interesting difference in where the cut between favoring and disfavoring contexts occurs. In Bislama, the cut looks like a rather simple presence versus absence in the prior clause (not present in the prior clause being the only context in which a free pronoun or NP is strongly favored). In Tamambo, the cut is sensitive to a more subtle interplay between presence in the prior clause and *continued* presence in the prior clause. A nonovert argument (whether subject or object) in the prior clause is likely to have occurred in some form or another in the penultimate clause as well. So the fact that a subject referent is favored to be null when it was a nonovert argument in the prior clause suggests that in Tamambo speakers may use null subjects with referents that are persistent and continuing topics across multiple clauses. I think we should take this effect seriously, because it emerges in Tamambo despite the very small data set over which the analysis of variation is being undertaken.

Strong transfer in this proposal involves cases where a feature in the grammar occurs variably in both varieties, there is overlap in the linguistic factors that act as significant constraints on the variable in both varieties, and there are parallels in the ranking of different factor groups in both varieties. That is, the factor group that is the strongest constraint in the model language is also the factor group exerting the strongest constraint on variation in the replica, the second strongest constraint in the model is the second strongest constraint in the replica, and so forth.

The data on reduction of the form of direct objects seem to illustrate an instance of this. In both Tamambo and Bislama, the strongest constraint on the variable was the grammatical role of the referent in the prior clause, and the second strongest constraint was whether the referent would be considered an alienable or inalienable possession. There were no other significant factors on either language.

However, as with the constraints on subjects, the devil proves to be in the details. In Tamambo, only objects that had been core arguments in the prior clause favor appearing as a clitic object in the current clause; in Bislama, the same kind of presence/absence constraint seems to be operating for objects as we saw for subjects. The referent's presence in any grammatical role in the prior clause strongly favors realization as a null object.

More interestingly, with respect to (in)alienable possession, we observed a complete reversal of the most strongly favoring factor. Referents that would be considered inalienable possessions (e.g., close kin, body parts) strongly favor clitics in Tamambo; in Bislama, these referents equally strongly *dis*favor phonetic reduction and tend to occur as overt nouns or pronouns.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

So, the differences between constraints within these factor groups mean that this variable fails to satisfy the criteria for a *calque*. Instead, it is an example of *strong transfer*. Under this very restrictive definition of a calque of a variable, calques will clearly be fairly rare. I would predict that they will only occur under conditions of lengthy or socially intense contact.

IMPLICATIONS OF THIS PROPOSAL

In previous work on the constraints on object deletion in Bislama, I suggested that there may be a very general principle of "transformation under transfer" (Meyerhoff, 2003). That proposal suggested that the process of transferring—or replicating more generally—patterns from the model to the replica language may be sufficiently cognitively demanding that speakers cannot replicate (or perhaps do not care if they replicate) the linguistic details of the model in the replica language. In that paper, I noted that the distinction between alienable and inalienable possession is marked (obligatorily) in possessive NPs in Bislama's substrate languages, whereas in Bislama itself the distinction never shows up in possessive NPs, instead it only probabilistically shows up in the distribution of overt direct object NPs.

The data discussed here seem to add weight and detail to the notion that transformation under transfer is a fairly general principle. We can see analogies in some of the nonvariable examples of replication that were discussed in the first half of the article. There, I noted that the inclusive/exclusive distinction in Bislama pronouns emerged through the creative transformation of English constituents into a pronoun system that replicated the substrate norms. Similarly, I noted that even in the case of a lexical calque, speakers are not blindly translating across systems. From *skyscraper* to *Wolkenkratzer* there is also a degree of transformation. I believe that we also saw transformation under transfer in the reversal of the favoring conditions for phonetically reduced direct object arguments in Bislama and Tamambo. The principle seems to operate at all levels of linguistic structure, even those so far below speakers' possible conscious awareness—the ranking of factors within a significant factor group that constrains a variable—which suggests that it is indeed a fundamental principle.

CONCLUSIONS

In this article, I have proposed integrating the tradition of investigating variation in creole languages with more recent models of comparative variationist sociolinguistics. The vast majority of creolists accept that in order to understand the joint influence of substrate and lexifier languages it is necessary to proceed by comparing features with some semantic equivalence in the language varieties that were in contact and in the creole outcome. Creolists often acknowledge the relevance of variation, especially to the early stages of the languages that are the objects of their enquiry. The principles proposed here outline a means by which these two fundamentals in creolistics can be enriched through the use of

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

variationist methods. It does this by proposing a way of particularizing how we define replication, transfer, and calquing to inherently variable domains of language.

In outlining a means for exploring the details of variation and for using this to inform fundamental questions in the field of language contact research, I have suggested that it is helpful to ask to what extent the variability in one language appears to be partially or fully replicated in another language. In doing so, we provide not only a basis for further enriching the long-standing connections between sociolinguistics and creolistics, but also provide a basis for linking research on creolistics and language contact to the growing field of comparative studies in variation.

NOTES

1. How a candidate is evaluated as "suitable" in cases like this is an interesting question. It is not at all clear to a non-German native speaker (like me) why Germans seem to find **Himmelkratzer* odd but not impossible, and **Wolkenschaber* completely bizarre or just plain funny. But that question is beyond the scope of this article.

2. When I first presented these principles in conference papers in 2007 (United Kingdom Language Variation and Change, Association for Linguistic Typology, and New Ways of Analyzing Variation), I used the terms *weak transfer, strong transfer*, and *calquing* only.

3. Spelt *Tamabo* in this publication. The spelling reflects the fact that nasalization of an intervocalic stop is entirely predictable, however *Tamambo* remains the more common spelling for now and is the one most people on Malo would recognize, so I will continue to use it in this article.

4. Attentive readers may notice some divergences in the constraint rankings in the Bislama data reported here and in the 2003 paper, which reported only on object deletion among speakers in Santo township. The source of that divergence (though interesting) is not the focus of the present article. I am grateful to the eagle eyes of Scott Schwenter for drawing my attention to the potential confusion arising from the discrepancies in constraint rankings of the Bislama for urban (Santo) and village (Malo) speakers.

5. Bislama examples identify speaker and tape recording (e.g., M-95-10, Malo 1995 tape 10); Tamambo examples identify *tandono*, teller, and date. Abbreviations in the glosses (which draw on Jauncey's [1997, 2002] analyses of Tamambo and my own of Bislama) are: AGR 'agreement', ASP 'aspect marker', CONT 'continuous aspect', FUT 'future irrealis', INDEP 'independent pronoun', IRR 'irrealis', PL 'plural', POSS 'possessive', SPEC 'specificity marker'.

6. One way of dealing with nonindependent factor groups like this is to compare separate analyses isolating the overlapping factor groups (Guy, 1993; Paolillo, 2002) and comparing the goodness of fit of the respective models.

7. At an earlier stage of the analysis of the Bislama data, coding for the animacy factor group included more fine-grained distinctions, that is, a distinction between first- and second-person versus third-person human referents, events/propositions, and abstract Ns. At this point, the factor group did emerge as a significant constraint, but as in the data shown in Table 3, the best fit (as measured by the lowest log likelihood, given the difference in the number of factors between the different runs [cf. Paolillo, 2002]) and ranking of factors constituted no logical animacy hierarchy. In the interests of providing a direct comparison with the smaller Tamambo data set, this detailed coding was collapsed, effectively imposing a canonical profile for animacy, and at this point the group failed to be identified as significant.

8. It is very likely that it will be useful to extend this to include social (nonlinguistic) constraints that are significant. However, other work in this general vein suggests nonlinguistic constraints are much less likely to transfer in part or whole (e.g., Buchstaller & D'Arcy, 2009; Meyerhoff & Niedzielski, 2003). As the current investigation has not looked at social factors, I have nothing further to say about this possible integration of linguistic and nonlinguistic factors at present.

REFERENCES

Blondeau, Hélène, & Nagy, Naomi. (2008). Subordinate clause marking in Montreal Anglophone French and English. In M. Meyerhoff & N. Nagy (eds.), *Social lives in language: Sociolinguistics and multilingual speech communities*. Amsterdam: John Benjamins. 273–313.

Britain, David. (2008). When is a change not a change? A case study on the dialect origins of New Zealand English. *Language Variation and Change* 20(2):187–223.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

Britain, David, & Peter Trudgill. (2005). New dialect formation and contact-induced reallocation: Three case studies from the English Fens. *International Journal of English Studies* 5(1):183–209.
Buchstaller, Isabelle. (2008). The localization of global linguistic variants. *English World-Wide* 29 (1):15–44.
Buchstaller, Isabelle, & D'Arcy, Alexandra. (2009). Localized globalization: A multi-local, multivariate investigation of quotative *be like*. *Journal of Sociolinguistics* 13(3):291–331.
Burridge, Kate. (1995). Evidence of grammaticalization in Pennsylvania German. In H. Andersen (ed.), *Historical linguistics 1993: Selected papers from the 11th International Conference on Historical Linguistics*. Amsterdam: John Benjamins. 59–75.
Camden, William G. (1979). Parallels in the structure and lexicon and syntax between New Hebrides Bislama and the South Santo language spoken at Tangoa. *Papers in Pidgin and Creole Linguistics, No. 2* (Pacific Linguistics, A–57). Canberra: Australian National University. 51–117.
Cotter, Colleen. (1994). Focus in Irish and English: Contrast and contact. *Berkeley Linguistics Society* 20:134–144.
Crowley, Terry. (1990). *From Beach-la-Mar to Bislama: The emergence of a national language*. Oxford: Oxford University Press.
_____. (2004). The transfer of fuzzy categories: Grammatical variability and animacy in Bislama. Paper presented at Montreal Dialogues: Workshop on SLA and Creole languages. University of Montreal.
Gundel, Jeanette K., Hedberg, Nancy, & Zacharski, Ron. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69(2):274–307.
Guy, Gregory R. (1993). The quantitative analysis of linguistic variation. In D. Preston (ed.), *American dialect research*. Amsterdam : John Benjamins. 223–241.
Heine, Bernd, & Kuteva, Tania. (2005). *Language contact and grammatical change*. Cambridge: Cambridge University Press.
Hock, Hans Heinrich. (1992). A note on English and modern Sanskrit. *World Englishes* 11(2/3): 163–171.
Jauncey, Dorothy. (1997). *A grammar of Tamambo*. Unpublished Ph.D. dissertation. The Australian National University.
_____. (2002). Tamabo. In J. Lynch, M. Ross, & T. Crowley (eds.), *The Oceanic languages*. Richmond, Surrey: Curzon. 608–625.
Jaegli, Osvaldo, & Safir, Ken. (eds.). (1989). *The null subject parameter*. Dordrecht: Kluwer.
Keesing, Roger M. (1988). *Melanesian Pidgin and the Oceanic substrate*. Stanford: Stanford University Press.
Kerswill, Paul, & Williams, Ann. (2005). New towns and koineization: Linguistic and social correlates. *Linguistics* 43(5):1023–1048.
King, Ruth. (2008). *Chiac* in context: Overview and evaluation of Acadie's *Joual*. In M. Meyerhoff & N. Nagy (eds.), *Social lives in language: Sociolinguistics and multilingual speech communities*. Amsterdam: John Benjamins. 137–178.
Koptjevskaya-Tamm, Maria, & Wälchli, Bernhard. (2001). The Circum-Baltic languages: An areal-typological approach. In Ö. Dahl & M. Koptjevskaya-Tamm (eds.), *Circum-Baltic languages: Typology and contact. Vol. II: Grammar and typology*. Amsterdam: John Benjamins. 615–750.
Le Page, R. B., & Tabouret-Keller, Andrée. (1985). *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
Matras, Yaron. (2000). Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism* 4(4):505–528.
Meyerhoff, Miriam. (1996). Transitive marking in contact Englishes. *Australian Journal of Linguistics* 16(1):57–80.
_____. (2000). The emergence of creole subject-verb agreement and the licensing of null subjects. *Language Variation and Change* 12(2):203–230.
_____. (2001). Another look at the typology of serial verb constructions: The grammaticalization of temporal relations in Bislama (Vanuatu). *Oceanic Linguistics* 40:247–268.
_____. (2003). Formal and cultural constraints on optional objects in Bislama. *Language Variation and Change* 14(3):323–346.
_____. (2008). Aspects of the language spoken at Matantas, Big Bay (Vanuatu). Unpublished manuscript, University of Edinburgh.
_____. (2009). Animacy in Bislama? Using quantitative methods to evaluate transfer of a substrate feature. In J. Stanford & D. Preston (eds.), *Variation in indigenous minority languages*. Amsterdam: John Benjamins. 369–396.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

Meyerhoff, Miriam, & Niedzielski, Nancy. (2003). The globalisation of vernacular variation. *Journal of Sociolinguistics* 7(4):534–555.
Meyerhoff, Miriam, & Walker, James A. (2007). The persistence of variation in individual grammars: Copula absence in "urban sojourners" and their stay-at-home peers, Bequia (St Vincent and the Grenadines). *Journal of Sociolinguistics* 11(3):346–366.
_____. (2009). Variation in the use of existentials on Bequia (St Vincent & the Grenadines). Paper presented at the Eighth Creolistics Workshop. Justus-Liebig-Universität, Gießen.
Mufwene, Salikoko S. (1996). The founder principle in creole genesis. *Diachronica* 13:83–134.
_____. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.
Naro, Anthony, & Scherre, Maria Marta Pereira. (2000). Variable concord in Portuguese: The situation in Brazil and Portugal. In J. McWhorter (ed.), *Language change and language contact in pidgins and creoles*. Amsterdam: John Benjamins. 235–255.
Poplack, Shana. (ed.) (2000). *The English history of African American English*. Oxford: Basil Blackwell.
Poplack, Shana, & Malvar, Elisabete. (2006). Modelling linguistic change: The past and the present of the future in Brazilian Portuguese. In F. Hinskens (ed.), *Language variation: European perspectives*. Amsterdam: John Benjamins. 169–199.
Poplack, Shana, & Tagliamonte, Sali. (2001). *African American English in the diaspora*. Oxford: Blackwell.
Roberts, Sarah J. (2000). Nativization and the genesis of Hawaiian Creole. In J. McWhorter (ed.), *Language change and language contact in pidgins and creoles*. Amsterdam: John Benjamins. 257–300.
Sankoff, David, Tagliamonte, Sali A., & Smith, E. (2005). Goldvarb X: A variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto.
Sankoff, Gillian. (1993). Focus in Tok Pisin. In F. Byrne & D. Winford (eds.), *Focus and grammatical relations in creole languages*. Amsterdam: John Benjamins. 117–140.
_____. (1994). An historical and evolutionary approach to variation in the Tok Pisin verb phrase. *Parasession on Variation and Linguistic Theory: Chicago Linguistic Society* 30:293–320.
Sankoff, Gillian, & Brown, Penelope. (1976). The origins of syntax in discourse: A case study of Tok Pisin relatives. *Language* 52(3):631–666.
Sankoff, Gillian, & Mazzie, Claudia. (1991). Determining noun phrases in Tok Pisin. *Journal of Pidgin and Creole Languages* 6:1–24.
Siegel, Jeff. (1999). Transfer constraints and substrate transfer in Melanesian Pidgin. *Journal of Pidgin and Creole Languages* 14:1–44.
Smith, Geoff P. (2002). *Growing up with Tok Pisin: Contact, creolization and change in Papua New Guinea's national language*. London: Battlebridge Publications.
Tagliamonte, Sali. (2002). Comparative sociolinguistics. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 729–763.
Thomason, Sarah Grey, & Kaufman, Terrence. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
Trudgill, Peter. (1986). *Dialects in contact*. Oxford: Basil Blackwell.
_____. (2004). *New dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
Turell, M. Teresa. (1998). The sociolinguistic situation of the British and US American communities in Spain. *Links and Letters* 5:193–200.
Winford, Donald. (2003). *An introduction to contact linguistics*. Oxford: Blackwell.

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press

APPENDIX

Paradigm of free and bound pronouns in Tamambo:

		Preverbal (bound) subject pronouns		
	Independent pronouns	Unmarked irrealis	Realis/nonfuture	Future
1s	iau	ku	ku	ku-mbo
2s	niho	o	o	o-mbo
3s	nia	a	mo	a-mbo
1p-incl	hinda	ka	ka	ka-mbo
1p-excl	kamam	ka	ka	ka-mbo
2p	kamim	no	no *(mo)*	ko-mbo
3p	nira	na	na *(mo)*	ka-mbo

Independent and bound pronominal forms in Tamambo. *Mo* variant in 2p and 3p realis is archaic.
Source: Jauncey (1997).

https://doi.org/10.1017/S0954394509990196 Published online by Cambridge University Press