Mixed languages: Re-examining the structural prototype

Yaron Matras

1. The Mixed Language prototype

A structural definition of Mixed Languages (MLs) can be expected to cover features which they typically share, and which distinguish them from other cases of contact-related change. Structural definitions proposed in the literature so far however have been to a large extent intuitive. The discussions in Bakker (1997), Bakker and Mous (1994), and Bakker and Muysken (1995) describe the ML prototype as showing a split between the source language of the 'grammar' and that of the 'lexicon', with variation within the class of 'function words'. What is implied (though not made entirely explicit) is thus a split between the source of bound or inflectional morphology, and that of unbound content words. This prototype is viewed by Bakker (1997) as an emblematic manifestation of hybrid identity, reflecting either (a) split allegiance to two parental communities, in cases where the parental generation's women and men have distinct origins, and the young generation forms a new ethnic group, or (b) maintenance of ethnic awareness in non-territorial minorities, often with peripatetic (itinerant or 'nomadic') cultures.

The grammar-lexicon split has been argued to reflect cognitive and communicative primitives. Bakker (1997) attributes the fact that mothers provide the 'grammar' language in the prototype mixed communities to their role in raising children. This could be understood to suggest a split, by source language, between the processing functions of language (grammar) and its referential functions (lexicon). Muysken (1997) brings 'relexification' in Media Lengua in connection with acculturation to dominant, urban Spanish-speaking society. This too can be understood to imply manipulation of the referential component of language for the purpose of identity-flagging. Myers-Scotton (1992, 1998) regards the grammar-lexicon split as a sedimentation of the constraints on mixed portions of speech, where a matrix and an embedded language provide system and content morphemes respectively. All these approaches have taken both 'grammar' and 'lexicon' largely as wholesale categories. They have also taken for granted that a borderline exists between MLs and other cases of contact-induced language change where grammar and

lexicon are affected, though not on a wholesale basis. The implicit structural statement about the ML prototype is therefore that it recruits its *entire* grammar and its *entire* lexicon from two different sources.

The prototype definition also includes aspects of the social circumstances of the emergence of MLs. There appears to be consensus in the literature (Thomason 1997a; Bakker 1997, 1998) that prototypical MLs emerge rapidly and in a situation of full bilingualism (which has earned the varieties in question the label 'Bilingual Mixtures', Thomason 1997a, 1997d), that they serve as markers of new ethnic identities, and that they evolve through a process of more-or-less conscious creation (but see criticism of the latter notion in Myers-Scotton 1998). An important further component of the ML prototype definition, one that is already highlighted by Thomason and Kaufman (1988), is that MLs constitute such extreme cases of contact-induced change that they confront the historical linguist with serious difficulties of genetic classification.

2. Problems with the prototype

The prototype ML entails a series of problems, deriving from the rather diverse nature of the small sample of well-described languages that have so far qualified as MLs in the literature:

a. Some languages do not adhere to the structural prototype. In Michif the division is not between grammar and lexicon, but roughly between noun phrase and verb phrase. Mednyj Aleut is an exception to the structural prototype in that finite verb inflection is Russian-derived, whereas nominal and nominalised bound morphology is Aleut. It seems that whatever structural constraints operate in these particular combinations of source languages (see Bakker 1997), they are powerful enough to override the suggested prototype formula.

b. The wholesale approach to 'grammar' and 'lexicon' leaves open the status of unbound grammatical operators, which may in principle pattern either with the grammaticiser or with the lexifier language. The methodology pursued by Jake (1994) and by Myers-Scotton (1992) suggests that function words that may have thematic status are more likely to pattern with content morphs. In Matras (2000) I proposed a tentative division between discourse-regulating function words, which are generally prone to borrowing; a class comprising existentials and negators, which are crucial to encoding meaning at the level of the entire proposition; and deictics, which are retrievable from situation-bound communication. These three properties – general contact susceptibility, propositional saliency, and situative retrievability – do not necessarily co-occur. General contact susceptibility is a product of the cognitive-processing functions of certain word classes in natural communication, typically involving a minority and a majority language (see discussion of the 'pragmatically dominant language' in Matras 1998). Propositional saliency is relevant to deliberate manipulations of the referential functions of language. Situative retrievability plays a role in conscious attempts to preserve impressions of an ancestral language for purposes of identity-flagging. What they can do is provide some clues toward an explanation of the source of certain classes of function words in different types of ML.

c. The view that the prototypical context of emergence necessarily involves situations of full bilingualism, puts varieties that are the products of language attrition, such as Ma'á or Para-Romani, in the periphery. It also excludes varieties that emerge through purposeful recruitment of foreign or manipulated (camouflaged) lexicon for use as a secret or in-group register. Both types resemble the ML prototype in that 'lexicon' from one source is combined with 'grammar' from another source. The function of these types – ancestral-language based ethnolects and secret languages – as symbols of ethnic separateness also fits the prototype. Bakker (1998) has argued against the inclusion of secret languages in the ML prototype on the grounds that active bilingualism is not involved in their formation, and that the source of their lexicon is typically diverse and includes camouflaged formations, and so it differs from the naturally acquired lexicon of prototypical MLs. On the other hand Bakker considers Para-Romani varieties as MLs, arguing that their structural profile testifies to an origin in a situation of bilingualism. The case of Ma'á has attracted considerable controversy, there being no attestation of full bilingualism at an earlier stage and not even conclusive evidence as to the identity of the variety that contributed the Cushitic material, while on the other hand the Cushitic component has been argued to be in decline throughout the documented lifetime of the language (Mous 1994; Thomason 1997b, Greenberg 1999).

d. The prototype assumes that MLs function as all-purpose community languages, and moreover that they are self-contained, coherent, and consistent linguistic systems, rather than on-the-spot and variable mixes at the utterance or discourse level. But this seems to be confirmed unambiguously only for Michif. In other cases of MLs, speakers have access to a non-mixed variety of one of the contributing languages, and in the case of Media Lengua in all likelihood even to both. To define the components of the ML in a case like Media Lengua or Mednyj Aleut means to be able to distinguish between stable system-components of (respectively) Spanish or Russian origin, and on-the-spot mixing involving these languages. There are more extreme cases where

the grammaticiser language is also the language of everyday interaction, and so the ML must be regarded as a mixed register of the grammaticiser language. These cases have been characterised by Smith (1994) as 'symbiotic' mixed languages. In his commentary on Matras (2000), Smith (2000) introduced a distinction between 'open' symbiotic MLs, which are registers that are employed at the discourse level, and 'closed' symbiotic MLs, which are confined to the level of individual utterances. The relation between the communicative functions of such systems and their structural profile is largely unexplored (but see discussion in Matras 2000). It is – in principle, at least – possible that Para-Romani varieties, for instance, represent the far edge of an historical continuum, ranging from an ML that arose in a situation of full bilingualism and could have been used for all-purpose communication (see Bakker 1998; Boretzky 1998), to a discourse-level register, and on to the mere occasional insertion of individual lexical items at the utterance level, as recently documented for present-day Para-Romani varieties (Leigh 1998; Matras 2002: ch. 10).

e. Finally, the prototype only provides an intuitive, not a systematic answer to the question what, in terms of the actual structural profile, makes MLs more mixed and less easily classifiable for purposes of historical and genetic reconstruction than cases of very heavy grammatical borrowing. Clearly, the greater the quantity in a language of structures that are inherited from the principal ancestor language, the easier it is for the historical linguist to reconstruct structural aspects of its evolution from that ancestor. But if quantity were the chief criterion for genetic classifiability, then surely one would expect conventional cases of contact-related change, and MLs, to form a continuum (see for this view Greenberg 1999, and partly also Thomason 1995). The assumption that the social processes that lead to the abrupt emergence of MLs correlate with distinct structural profiles, implies that MLs display patterns of mixing that are not generally found in cases where contact-related structural change is a gradual development. In other words, MLs are considered exempt from the scale of borrowability and the tentative constraints or at least hierarchies that it entails. This view is captured by the distinct labels that have been assigned to the varieties – Mixed Languages, or Bilingual Mixtures.

3. General outlook: The layered nature of mixed languages

A model of MLs should allow to reconcile a revised prototype with the structural and functional diversity of MLs. My aim in this chapter is to revise the notion of the ML structural prototype by examining a small sample of MLs for which documentation is more readily available. I first review some of the

observations on what is rarely borrowable in grammar and structure, drawing on the example of a language with heavy borrowing, in order to help establish just where and how the structures displayed by MLs are unique. I then proceed to examine the sources of individual classes of grammatical functions in the sample languages. My conclusions are the following:

a. MLs show remarkable consistency in the source language that provides the finite verb inflection, for which I use in a pre-theoretical sense the term 'INFL-language'. The INFL-language also supplies the rules for word order in the verb phrase, and typically also the typology and structures of clause combining. The notion of a 'grammar' or 'matrix' language is therefore better defined in terms of an INFL-language that is responsible for the anchoring of predications and so for the overall processing of propositional contents. The absolute consistency of the INFL-language allows to consider structures that do not derive from the INFL-language as 'borrowings', 'insertions', or otherwise as instances of structural incorporation from a contact language. Historical continuity of the INFL-language from the parental generation to the first generation of an ML's native speakers, implies that the system in which predications are anchored and propositions are processed is continued; thus, in Media Lengua there is continuation of Quechua as INFL-language. Discontinuity of the INFL-language implies completion of the most significant step toward language shift; thus, Mednyj Aleut shows shift of the INFL-language from Aleut to Russian.

b. The exceptional feature of MLs, is that they display a split between the INFL-language, and the source of most unbound, potentially autonomous content words, especially nouns. This is the essence of the grammar–lexicon split in the prototype. Historically, there are two ways in which the prototype may emerge. Both involve manipulation of the referential or content component of language (see Matras 2000). The first is 'lexical re-orientation': speakers hold on to their ancestral INFL-language, but use lexical insertions from an L2 (for purposes of camouflaging meaning, or flagging acculturation). The second is 'selective replication': speakers have shifted to a new INFL-language (cf. Myers-Scotton's 'Matrix Language Turnover', 1998), but maintain impressions of the ancestral language, for similar purposes (camouflaging meaning, or defying acculturation pressure). Both processes are exceptional, in that they involve a kind of natural (rather than institutionalised) language design by which a community of speakers uses language consciously and deliberately in order to project its image of itself onto the structural composition of its speech. But while identity-flagging through language mixing is common in bilingual communities, and can even be argued to lead to long-term language change both at the idiolect level (Maschler 1998) and at the commu-

nity level (Auer 1999), it does not always lead to the wholesale substitution of the content-lexical component, as it does in MLs.

c. Lexical re-orientation and selective replication target the lexicon in the first instance. But the split between INFL-language and the source of the lexicon appears also to grant speakers a licence to borrow grammatical structures that are not normally borrowed even in situations of very intense and prolonged contact. This clash between the INFL-language and the source of grammatical categories that are, at least in tendency, 'loan-proof', is an additional distinctive feature of the ML prototype that stands out regardless of the structural variation among individual MLs.

d. The impression that the split between INFL and the content-lexicon allows speakers to 'break the rules' on grammatical borrowing in other categories as well has been conveyed by Bakker (1997 and earlier; also paraphrased in Golovko 1994). It must be noted however that MLs that serve as all-purpose languages for everyday communication, and that are, as minority languages, in contact with a dominant majority language, will also be subjected to the conventional patterns of contact influence. To some extent, it is possible to make function-related predictions about the composition of some grammatical classes (cf. Matras 2000). Thus, conjunctions and discourse-markers will pattern with the INFL-language, unless they are borrowed from a dominant majority language (for instance English markers in Michif). Existentials and negators have high referential value (propositional saliency) and so they are likely to pattern against the INFL-language especially where the ML serves as a secret register. Demonstratives are high on the deictic-situative scale and can only pattern against the INFL-language if a situation of bilingualism existed in the earlier history of the ML.

e. Apart from such function-related constraints, the structural profile of an ML will depend on the typological similarities and clashes between the contributing languages. The emergence of the structural profile of an ML is therefore best described in terms of the density of several types of language contact phenomena, or Layers.

The point of departure is a consistent *INFL-language* in which the predication is processed. The historical inner core or Layer 1 of the ML prototype consists of the wholesale incorporation of autonomous (unbound) lexicon from a *lexifier language*.

Layer 2 is an extension to the core: grammatical components that are rarely borrowable pattern with the lexifier language, since they are considered crucial to meaning (existentials, negators, personal pronouns) or representative of communicative interaction in the lexifier language (deictics, lower numerals). The core, or Layer 1, and the extension to the core, or Layer 2, are the only

components that are constitutive of the ML prototype. Both are found only in MLs and so they form part of the ML definition. But Layer 1 is constitutive of all MLs, while Layer 2 depends on the types, the functions, and the emergence circumstances of individual MLs, and so it is subject to greater variability.

Layer 3 represents the adaptation effects associated with the incorporation of a foreign lexicon (case markers, adpositions, attributive deictics, agreement markers) and consists similarly of elements that are less likely to be borrowed, though they are more commonly attested in languages with heavy borrowings that are not MLs than the phenomena of Layer 2. The outcome can be a convergent re-structuring of a category, or the wholesale incorporation of a class of items (such as prepositions or attributive demonstratives) from the lexifier language. I have referred to the latter kind of process – the wholesale incorporation of a class of items – as 'fusion', while the effects of convergence that does not involve the incorporation of actual forms can be captured through Ross's (1996) term 'metatypy'.

Layer 4, finally, is the outer layer where an ML, like other languages in contact, may incorporate structures and even entire classes of grammatical items from its contact language. It is here that more conventional processes are attested, such as the borrowing of phasal adverbs, conjunctions, focus particles, modality markers, and more.

4. Constraints on borrowing and language mixing

As language contact acquires a more prominent position in approaches to language change, the rich amount of evidence of contact-induced change has prompted scepticism with regard to the formulation of any absolute constraints on structural borrowing (Thomason and Kaufman 1988; Harris and Campbell 1995; Thomason 2001). Nonetheless, tentative generalisations about the relative likelihood of borrowing have been made in the literature. It is generally accepted that lexical items are more frequently borrowed than grammatical items, that nouns are more frequently borrowed than verbs or adjectives, that unbound elements are more easily borrowed than bound elements, and that derivational morphology is more easily borrowed than inflectional morphology (cf. Haugen 1950; Weinreich 1953; Moravcsik 1978; Muysken 1981; Thomason and Kaufman 1988; Romaine 1989; Campbell 1993). The overall picture from a structural perspective thus points to the relative autonomy of the word (compared with the tightly-integrated status of the inflectional morph) as a factor that facilitates borrowing (on 'paradigmaticity' as an inhibiting factor cf. van Hout and Muysken 1994).

Apart from the issue of the structural properties of borrowable elements, three additional observations are noteworthy. First, the arrangement of categories, i.e. the functional scope of individual morphs such as TAM-markers, their relative position, and syntactic properties such as word order and the typology of clause linking, may undergo a process of convergence without the actual transfer of phonological forms from one language for replication in another. Convergence of this kind on a wide scale has been argued to be constitutive of linguistic areas (Masica 1976; Joseph 1983; Ross 1996), and typical of language change in bilingual communities (Silva-Corvalán 1994). Next, the susceptibility of grammatical categories to contact-related change has been shown to correlate with their degree of discourse-prominence, with elements that operate at the discourse level at the top of the borrowing-likelihood hierarchy, followed by clause-linking devices, and word-level phenomena at the end of the cline (Stolz and Stolz 1996; Matras 1996). Finally, borrowing has been shown to involve in some instances not just individual items, but entire classes, leading to a wholesale merger of a category of elements in the two languages (Salmons 1990), or 'fusion' (Matras 1998). Cases of fusion involve such classes as discourse operators, focus particles and phasal adverbs (Matras 1998), but also expressions of aktionsart such as Slavic aspect marking in dialects of Romani, or Arabic prepositions in Domari (Matras 1999). Fusion seems to be of particular interest to the ML discussion due to the categorial compartmentalisation by source language which MLs typically display.

Of special interest to the status of MLs are tentative generalizations on linguistic structures that may appear to be 'loan proof' in conventional situations of contact and contact-related structural change ('conventional' situations being those in which the outcome of the change does not lead to ambiguity with regard to the genetic classification of a language). The most explicit attempt at formulating a reliable indicator for genetic classification is the Swadesh wordlist of core vocabulary. Although many of the items on the list are clearly borrowable (English having borrowed up to 30 per cent of its Swadesh-list vocabulary), wholesale borrowing of the list inventory would amount by the list's working definition to a significant disruption of historical-genetic continuity. On the grammar side, although no absolute constraints can be formulated and the possibility of counter-examples must always be left open, there are nevertheless a number of categories for which borrowing of actual phonological forms (rather than the modelling of a category on a contact language using inherited forms) is only rarely attested. They include definite and indefinite articles, bound person markers on the verb (concord), bound possessive markers on nouns, bound tense and aspect markers (including infinitival and possibly other converb markers) on the verb, bound case affixes, sentential negators (expressing negation of the finite indicative predicate), personal pronouns, demonstratives, existentials (copula), the lexical verb which expresses the possessive notion 'to have', place deictics, the basic interrogatives 'what' and 'who', numerals under '5', as well as adpositions that cover basic local relations ('in', 'at', 'out of').

While individual exceptions are attested for almost all of the items on this list, they generally involve single elements from a class, rather than the entire class itself. Two generalisations can be made on the conditions under which the constraints of the tentative 'loan-proof' grammar are relaxed somewhat. The first involves borrowing in interdialectal contact. The second involves replication of the borrowed elements only with borrowed vocabulary, with no productive diffusion into inherited vocabulary. Thus, Romani dialects in contact with Russian and Turkish replicate the respective verb inflection (including both concord and TAM markers) with verbs from the respective languages (Rusakov and Abramenko 1998; Igla 1996), Moroccan Arabic tends to incorporate French nouns together with their definite or indefinite articles (Bentahila and Davies 1995). Such patterns of grammatical compartmentalisation are clearly connected to the availability of the source language as an active system (fossilised in the case of Agia Varvara Romani, Igla 1996).

Studies of codeswitching suggest that some constraints on structural mixing apply even where speakers have full active access to the contact language synchronically. Perhaps the most general statements on structural mixing are Poplack's (1980) 'free morpheme constraint' and Myers-Scotton's (1993) 'system morpheme hypothesis', both of which predict constraints on the transfer of individual grammatical morphemes. Evaluations of the proportion of individual word classes among instances of codeswitching reveal that switching is on the whole rather rare around personal pronouns, definite articles, and adpositions (Pfaff 1979; Poplack 1980; Berk-Seligson 1986; Backus 1992; Bentahila and Davies 1995).

The broader picture therefore confirms that there are classes of elements in grammar that are less likely to be transferred among languages, both in situations of synchronic mixing and in cases of diachronic change involving contact. The density with which these general tendencies are violated in MLs merits particular attention. But do MLs differ in principle from languages with 'heavy borrowing' (in the sense of Thomason and Kaufman 1988)? As a point of comparison I consider here Domari, an Indic language spoken by peripatetic communities of the Near East, best described for the variety of a community of former metalworkers in Jerusalem (Macalister 1914; Matras 1999). Domari has been in contact with Arabic since medieval times. Oral

tradition suggests that the Dom may have arrived in Arabic-speaking regions in the thirteenth century, though no historical records are available to confirm this. Samples of Palestinian Domari date back to the early nineteenth century. A comparison between the material collected in the early twentieth century (Macalister 1914) and present-day Domari (Matras 1999) shows the decline of Indic-derived numerals, and a spread of Arabic inflection with Arabic auxiliaries (used earlier in an uninflected form). Today's Domari-speaking community in Jerusalem comprises only several dozen people, the younger generation having shifted to Arabic. Domari has thus retreated even as a language of the family. Table 1 offers an overview of Arabic influence on the language.

Domari employs exclusively Arabic material for all clause-linking devices, modals (with the exception of 'can') and the aspectual auxiliary of habituality. The use of the Arabic resumptive pronoun, and the bound agreement inflection of the resumptive pronoun and complementiser, can be classified as part of the Arabic clause-linking structures. Modal and aspectual auxiliaries retain Arabic verb inflection, including person concord and TAM, as well as negation. All unbound prepositions are Arabic, as are most person-inflected prepositions (with the exception of 'on' and 'for'). The Domari enclitic copula and the existential verb appear alongside the Arabic copula *kān* 'to be' and existential *ṣār* 'to become' in non-present indicative forms; this can partly be explained through the overlap of functions with the aspectual and modal auxiliaries *kān* and *ṣār* respectively. The point to be made in connection with Domari is that replication from the contact language of bound morphology (person, TAM, agreement) and anaphora does occur, but it is linked and confined to clause combining strategies or auxiliary modification, and in any case there is no diffusion of these structures into the inherited, non-Arabic component of the language. There is no paradigm borrowing of personal pronouns or of demonstratives.

5. The sample

To what extent do MLs clash with general tendencies to contain the diffusion of borrowed material within certain limits? I consider here a small sample of languages that have been referred to as 'mixed' in the literature, in relation to the assumed prototype described above. Discussions of Ma'á (Mous 1994; Thomason 1997b), Michif (Bakker 1994; Bakker and Papen 1997), Media Lengua (Muysken 1994, 1997), Mednyj Aleut (Golovko 1994; Thomason 1997c), and Para-Romani varieties (Boretzky and Igla 1994) have been included in previous collections on MLs. For Para-Romani, I draw on avail-

Table 1. Distribution of grammatical classes by source in Domari

Swadesh list lexicon	Ca. 50% pre-Arabic	Ca. 50% Arabic
Numerals below 5	Indic	(3–4 also Arabic)
Numerals 10, 100	Indic	
Other numerals above 5		Arabic
Bound case markers	Indic	
Indefinite article	Indic	
Interr. 'what', 'who', 'where'	Indic	
Interr. 'how much'		Arabic
Person concord (verb)	Indic	
Possessive markers (noun)	Indic	
Bound plural markers	Indic	Arabic
Bound tense markers	Indic	
Bound aspect markers	Indic	
Bound modality markers	Indic	
Bound converb markers	Indic	
Lexical verb negation	Indic	
Non-verbal predication marker	Indic	
Non-verbal predication negation		Arabic
Modal 'can'	Indic	
'Can' inflection and negation	Indic	
Other modals		Arabic
Other modal inflection and negation		Arabic
Aspectual auxiliaries		Arabic
Auxiliary inflection		Arabic
Auxiliary negation		Arabic
Existential present	Indic	
–negation	Indic	
Existential past/subjunctive	Indic	Arabic
–negation	Indic	Arabic
Personal pronouns	Indic	
Demonstratives	Indic	
Resumptive pronoun		Arabic
Resumptive pronoun agreement		Arabic
Bound adpositions (person marked)	Indic	Arabic
Unbound adpositions		Arabic
Conjunctions		Arabic
Complementiser agreement		Arabic
Phasal adverbs		Arabic
Focus particles		Arabic

able sources for the varieties of England, Spain, Germany, and Scandinavia, as well as on recent fieldwork in northern England. Lekoudesch (and related names) is the secret, in-group and trade language employed by Jewish cattle-traders in Germany until the 1930s. Its 'grammar' language is dialectal German, its 'lexicon' derives from non-spoken Ashkenazic Hebrew. The material used here was collected among non-Jewish farmers who had acquired the register (Matras 1991; Klepsch 1996), but historical records of the variety exist in the form of glossaries, going back to the seventeenth century. Jenisch is the term employed for a variety of secret and in-group lexicons used by peripatetic groups (sedentary with itinerant occupations) in south-western Germany and Switzerland. The material considered here (some 150 items from the village of Unterdeufstetten, on the Württemberg/Bavarian border; see also Dürr 1961) is composed of Rotwelsch vocabulary, much of it attested in earlier sources as far back as the fourteenth century, with some 30 per cent Romani-derived vocabulary, some 15 per cent Hebrew-derived items, and additional items of French origin. This kind of mixture is rather typical of the Jenisch varieties attested in southern and central Germany, Switzerland, and (under the name 'Bargoens') in the Netherlands.

The sample is small, but it is representative of the functional and structural diversity that one finds in cases of languages intuitively classified as 'mixed'. Three of the languages – Michif, Media Lengua, Mednyj Aleut – have been described as all-purpose languages that are sufficiently distinct from either of their respective source languages to be treated as separate systems. Ma'á differs somewhat, in that it can be viewed as a register of the Bantu language Pare, and so as symbiotic. Switches into Ma'á may however cover entire stretches of discourse. Present-day Para-Romani, Lekoudesch, and Jenisch on the other hand all consist of lexical inventories of several hundred words each of which are employed at the level of individual utterances (see Matras 2000 on the discourse position of mixed utterances).

The degree of bilingualism during the emergence of the languages is disputable in some cases. For Ma'á, the contact situation that led to the ML is believed to have involved a Cushitic language, and a Bantu language, though the identity of neither is established, and the principal Bantu component in today's Ma'á is supplied by Pare. The puzzling feature about Ma'á is the considerable lexicon of unknown origin, which resembles the recruitment of camouflaged lexicon in secret languages. The lexical component of Para-Romani is on the whole coherent and so the varieties appear to have emerged in a situation of bilingualism in Romani and a variety of the respective majority language. But there are some tendencies to incorporate lexicon from neighbouring secret languages. The genesis of Lekoudesch involves, exceptionally, literary bilingualism. Ashkenazic Hebrew was not spoken, but used in Jewish communities as a literary language. Through proficiency in Hebrew, speakers had access to a complete lexicon and grammar, protected from language attrition. The creative processes of compounding and derivation within Lekoudesch however testify that the Hebrew-derived lexicon acquired an independent status. As transmitted by non-Jewish speakers, at least, Lekoudesch also seems to include a number of items of non-Hebrew origin, mainly French, that are attested in other secret languages of the region. Jenisch is not connected to active bilingualism at the community level, though the absorption of ethnic minorities and immigrants within the Jenisch population could have been the source for some vocabulary influence. The principal sources of the Jenisch vocabulary are camouflaged items, and the ethnic secret vocabularies of neighbouring populations (Romani-, Hebrew-, and French-based).

Of the sample languages, then, only Jenisch, and to some extent Ma'á, can be said to lack a principal lexifier language. Para-Romani, Mednyj Aleut, and apparently also Ma'á can be considered products of language shift and partial language death. Only speakers of Media Lengua, some speakers of Michif, and users of Lekoudesch who were literate in Hebrew can be said to have active access to both source languages at the point of the creation of the mixed varieties. While all languages are markers of ethnic identity, the use of Lekoudesch, Jenisch, and to some extent Para-Romani is also associated with certain trade contexts.

6. INFL-language

All of the sample languages show an overwhelming tendency toward consistency in the language that provides the structures employed to process the predication. I adopt, in a pre-theoretical sense, the term *INFL-language* to denote the source of the grammatical structures involved in anchoring the predication. The leftmost position in Table 2 shows the source languages of the lexical root of the verb, which is typically the lexifier language. Michif is the only exception, showing more Cree-derived verbs and only some verbs from the lexifier language, French. Bakker (1997) argues that Cree does not have unbound lexical verbs and Cree verb morphology cannot be isolated from its stem, and therefore it cannot be combined with French roots. The boundness constraint therefore prevents lexical re-orientation from affecting the verb. Aleut has borrowed Russian vocabulary, and so some of the Mednyj Aleut verb roots are Russian-derived. In the secret languages Para-Romani, Lekoudesch and Jenisch, the presence of material from the lexifier language is al-

Table 2. The source of predication-anchoring grammar in the sample

Language	Verb-root	Concord	TAM	Converb	Infinitive	V. deriv.	Modal
Ma'á	Cush etc	Pare	Pare		Pare	Pare	Cush, etc.
Michif	Cree/(Fr)	Cree	Cree	Cree		Cree	Cree/Fr
MedLeng	Span	Quech	Quech	Quech	Quech	Quech	Span
MedAleut	Aleut/(Rus)	Rus	Rus	Aleut	Rus	Aleut	Rus
ParaRom	Rom	Non-R	Non-R		Non-R	Non-R	(Rom)
Lekoud	Hebr	Germ	Germ		Germ	Germ	Germ
Jenisch	various	Germ	Germ		Germ	Germ	Germ

ways a matter of choice at the level of the individual utterance, and the appearance on the matrix reflects the availability of a structure, rather than its obligatory use.

Most consistent across the sample is the marking of person-concord on the verb, and of bound tense-aspect-modality, as well as infinitive marking. Only Mednyj Aleut has converbs that pattern with the lexifier language Aleut, in line with the overall retention of nominal inflection from Aleut, as well as verb derivation. Modal verbs display tendencies to pattern with the lexifier language, i.e. they behave like lexical verbs. Exceptions are Mednyj Aleut, where Russian modal verbs replace Aleut modal affixation, and Lekoudesch and Jenisch, which have no historical access to an actively spoken lexifier language and where the motivation is most clearly to conceal meaning rather than to maintain impressions of an ancestral language, or to flag acculturation.

The INFL-language is, throughout the sample, also the language that supplies the word order rules for the verb phrase. It is overwhelmingly also the source of clause combining devices (Table 3). Exceptions to this are apparently the survival of a coordinating conjunction *neri* 'or' and of a conditional conjunction *kái*, both of unknown origin in Ma'á, alongside Pare equivalents; the use of French-derived conjunctions *pi, ou-ben-donc, a-cause, apre-k-* in Michif and of Spanish conjunctions such as *i, ki, porke* in Media Lengua; and the use of Aleut converbs for temporal and purpose clauses in Mednyj Aleut. For Media Lengua, Spanish infiltration of the clause combining system can be explained through the continuing presence of Spanish as the surrounding majority language. Possibly, the presence of French conjunctions in Michif might reflect the prestige of French at an earlier stage.

The consistency of the source language in anchoring and combining predications allows to identify a clear base-language or 'matrix-language' in the evolution of each ML. The correlation among categories that are relevant to

Table 3. Sources of clause combining devices in the sample languages

Language	INFL	Coordination	Complement	Conditional	Temporal	Relative
Ma'á	Pare	Pare/?		?		Pare
Michif	Cree	Cree/Fr	Cree	Cree	Cree/Fr	Cree/Fr
MedLeng*	Quech	Quech/Span	Quech/Span	Span/Quech	Quech	Quech
MedAleut	Rus	Rus	Rus	Rus	Aleut/Rus	Rus
ParaRom	Non-R	Non-R	Non-R	Non-R	Non-R	Non-R
Lekoud	Germ	Germ	Germ	Germ	Germ	Germ
Jenisch	Germ	Germ	Germ	Germ	Germ	Germ

Language	INFL	Causal	Purpose	Discourse markers	Majority language
Ma'á	Pare	Pare	Pare	Pare/Swahili	Pare/Swahili
Michif	Cree	Fr/ Cree	Cree	English/Fr/Cree	English
MedLeng*	Quech	Quech/Span	Quech	Span	Span
MedAleut	Rus	Rus	Aleut	Rus	Rus
ParaRom	Non-R	Non-R	Non-R	Non-R	Non-R
Lekoud	Germ	Germ	Germ	Germ	Germ
Jenisch	Germ	Germ	Germ	Germ	Germ

the verb, the correlation with clause combining patterns, and the absolute consistency in the marking of concord and TAM, all suggest that it is not justified to view MLs as cases of borrowing of verb inflection (cf. Greenberg 1999 on Ma'á). Due to the role of the INFL-language as the carrier of the predication and its mental anchoring, one cannot in my view regard the 'grammar' and 'lexifier' languages as having equivalent hierarchical status in the evolution of MLs, either. Rather, the INFL-language is the base into which lexifier language material is incorporated. In some MLs, the INFL-language reflects historical continuity of the language of the ancestral generation (Media Lengua), or of the language of the mothers (Michif). This is also the case in the creation of secret registers (Lekoudesch, Jenisch). In other cases, the INFL-language represents a language shift that has taken place or is ongoing (Ma'á, Mednyj Aleut, Para-Romani).

The extraordinary feature of MLs is the wholesale incorporation of lexical material from a source that is not the INFL-language. The only partial exception to this generalization are Cree verbs, as explained above (see Table 4).

An indication of the lexical composition of MLs is provided by the values for the combined Swadesh lists (100 and 200 word list, giving altogether

Table 4. Sources of nouns and verbs

Language	INFL	Lexical noun	Verb root
Ma'á	Pare	Cush, etc.	Cush, etc.
Michif	Cree	Fr	Cree/(Fr)
MedLeng	Quech	Span	Span
MedAleut	Russ	Aleut	Aleut
ParaRom	Non-R	Rom	Rom
Lekoud	Germ	Hebr	Hebr
Jenisch	Germ	various	various

207 different items). For Ma'á, up to 80 per cent of the attested basic vocabulary does not derive from the INFL-language. Possible Cushitic-derived items account for some 25–40 per cent of the vocabulary, Bantu elements for some 17 per cent, Maasai loans constitute up to 5 per cent, with the remainder of unknown origin. For Media Lengua, Spanish, the lexifier languages, provides some 85 per cent of the list vocabulary. Similarly, 90–95 per cent of Mednyj Aleut entries are Aleut-derived. For Michif the figure is balanced by comparison, with only around 48 per cent from the lexifier language French, and some 52 per cent from Cree, due to the constraints on the replacement of Cree verbs. According to Boretzky (1998), Para-Romani varieties show between 80–90 per cent Romani-derived vocabulary on the shorter, 100-item list. The figure for Spanish Para-Romani (Caló) drops to around 70 per cent for the longer, 207-item list.

7. Extended core grammatical classes

The split between INFL-language and the source of most of the basic vocabulary is accompanied by a licence to incorporate classes of grammatical items that are otherwise rarely borrowable. I refer to them as the extended core of the ML prototype. Table 5 displays the sources for a number of tentatively 'loan-proof' grammatical categories in the sample languages: the copula, sentential negation (of the finite verb), the verb in possessive constructions (*to have*), and the genitive-possessive construction (*X of Y*):

Among plain (all-purpose) MLs where the INFL-language is a continuation of the historical INFL-language, namely Michif and Media Lengua, the tendency is for the copula to pattern with the INFL-language. In Michif, the copula verb *ašte-* and copula infix *-iwi-* as well as the zero-copula, all from

Table 5. Sources for some grammatical categories

Language	INFL	Copula	Negation	'To have'	Gen. poss.
Ma'á	Pare	Cush/Pare	Pare	Cush/Pare	Pare?
Michif	Cree	Cree/Fr	Cree/Fr	Cree	Fr/Cree
MedLeng	Quech	Quech/Span	Quech/Span	Span	Quech?
MedAleut	Rus	Aleut	Rus	Aleut	Aleut
ParaRom	Non-R	Rom/(Non-R)	Rom/Non-R	Non-R	Non-R
Lekoud	Germ	Hebr/(Germ)	Hebr	Germ	Germ
Jenisch	Germ	Hebr	Rom	Germ	Germ

Cree, are continued, but they are supplemented by French-derived presentative and attributive particles (*si* < *c'est* and *ili* < *il est*). In Media Lengua, the Quechua copula *ga-* is continued, but supplemented by Spanish-derived *abi* 'there is'. Both French copulas in Michif, and Spanish presentatives in Media Lengua, can therefore be considered stereotypical replications. Where a shift in the historical INFL-language has taken place (selective replication), there is a clearer tendency toward replication of the copula from the lexifier language. A Cushitic-derived existential copula *he-lo* appears to be attested in earlier forms of Ma'á. In Mednyj Aleut, the Aleut copula *ū-* is replicated. Para-Romani varieties differ; in some sources for British and Spanish Para-Romani, the Romani copula in *sh-/sin-* is continued. In Scandinavian Para-Romani, new existentials are created from the lexical verbs *besj-* 'to sit' and *asj-* 'to stay', as well as from the modal *honke* < *hom-te* 'must'. Similarly, both Lekoudesch and Jenisch employ as a copula the Hebrew-derived *scheff-* < 'to sit'. There is thus a tendency in languages employed as secret registers to disguise the existential constructions, which is in line with the use of the ML to conceal propositional content. Also in line with the same function is the replication of lexifier-language material for negators, which are high on the propositional saliency scale, but not for genitive possessive constructions, which are grammatical. As for 'to have', Para-Romani, Lekoudesch and Jenisch can be said to lack a model for replication in the lexifier language. In Michif and Media Lengua, 'to have' patterns as a lexical verb, with Cree *aya-* and Spanish *tini-* respectively. In Mednyj Aleut, the Aleut copula supplies the verb, but the possessor follows the Russian model (*u minja . . . ū-*).

Overall, the copula tends to pattern with the INFL-language where the INFL-language is historically continued (lexical re-orientation), and with the lexifier language when there has been a shift in INFL-language. Negation patterns with the INFL-language in plain and discourse-level MLs, although

lexifier-language negators may also infiltrate, and with the lexifier language in secret, utterance-level MLs. The verb 'to have' is treated as a lexical verb in plain and discourse-level MLs, but as a grammatical construction in the secret MLs of the sample. The genitive-possessive construction appears to pattern with noun phrase grammar, discussed below.

I now turn to another segment of tentatively 'loan-proof' grammar – deictics and pronominals (Table 6). A simplified generalization might be: deictics and pronominals tend to pattern with the language that supplies nouns, provided that the ML arose in a situation of active bilingualism and speakers had access to actual communicative interaction in an all-purpose language. This excludes Lekoudesch and Jenisch, though it does not exclude secret languages in principle, since many are known to show manipulated or camouflaged forms of personal pronouns (see Smith 1998). The exception that stands out is Michif, which derives its nouns from French, but its deictics from Cree. Some fossilised French pronouns are used in Michif with French-derived verbs, usually in stereotypical, fixed expressions. Possessive pronouns in Michif pattern with the noun phrase, that is, they are derived from French, with the exception of some cases of inalienable possession. In Media Lengua, plural personal pronouns often combine a Spanish stem and a Quechua plural marker, modelled on the Quechua system of pronouns (*miu* 'I', *miu-kuna* 'we'). In Mednyj Aleut, Russian pronouns accompany past-tense verbs where person-agreement is lacking, and so they can be treated as extension of the Russian-derived system of person concord and so within the constraint of INFL-language consistency. The tendency is thus for cases of selective replication to replicate lexifier-language deictics. This tendency is also found in Media Lengua, though the Quechua paradigmatic system is partly maintained, resulting in hybrid formations.

Table 6. Sources of deictics and pronominals in the sample languages

Language	INFL	Demonstratives	Pers. pron	Poss. pron	Interrogatives
Ma'á	Pare	Cushitic	Cushitic	Cushitic	Pare?
Michif	Cree	Cree	Cree/(Fr)	Fr/(Cree)	Cree
MedLeng	Quech	Span	Span/Quech	Span	Span
MedAleut	Rus	Aleut	Aleut/Rus	Aleut	
ParaRom	Non-R	Rom/Non-R	Rom	Rom	Rom/Non-R
Lekoud	Germ	Germ	Germ	Germ	Germ
Jenisch	Germ	Germ	Germ	Germ	Germ

8. Noun phrase grammar

Finally, I turn to noun phrase grammar. Having established the consistency of the INFL-language in the sample MLs, it is noteworthy that noun phrase grammar, including inflection grammar, does not always pattern with the INFL-language (Table 7). Cree-derived nominal inflection in Michif refers to the continuity of the obviative suffix. In Media Lengua and, to the extent that nominal inflection is at all productive in German dialects, also in Lekoudesch and Jenisch, as well as in Basque, Turkish, and German Para-Romani, nominal inflection patterns with the INFL-language.

Several of the more outstanding features in Table 7 are found in other cases of heavy borrowing. It is not uncommon for borrowed nouns to be accompanied by borrowed plural markers, or even for plural markers to be adopted productively. Only low numerals are more commonly retained in the cases of selective replication (Ma'á, Mednyj Aleut, Para-Romani), which is also the case in Domari. Massive borrowing of prepositions is also found in Domari. Exceptional is the replication of the definite (and indefinite) articles of Spanish and French nouns in Media Lengua and Michif respectively, the preservation of French adjectival agreement and noun phrase word order (not included in the Table) in Michif, and the replication of Aleut case inflection in Mednyj Aleut. On the whole, then, most features are not extraordinary, in the sense that they are exclusive to MLs. Those that are, are only found in some of the sample languages. I therefore consider the features displayed here to form part of the 'accommodation layer' for nominals from the lexifier-language; these may or may not pattern with the lexifier language, and the important fea-

Table 7. Sources of noun phrase grammar

Language	INFL	Lexical noun	Plural	Definite article	Case infl.	Num.	Adposition
Ma'á	Pare	various	Pare			Cush/Pare	Cush, etc.
Michif	Cree	Fr	Fr	Fr	(Cree)	Fr	Cree/Fr
MedLeng	Quech	Span	Quech	Span	Quech	Span	Span
MedAleut	Russ	Aleut	Aleut		Aleut	Aleut/Rus	Aleut
ParaRom	Non-R	Rom	Non-R/ Rom	Non-R	(Non-R)	Rom/ Non-R	Non-R/ Rom
Lekoud	Germ	Hebr	Hebr/ Germ	Germ	(Germ)	Hebr	Germ
Jenisch	Germ	various	Germ	Germ	(Germ)	Germ	Germ

ture of MLs is that their processes of evolution license non-conventional accommodation.

9. Conclusions

What do the sample languages have in common, and how does this differ from other cases of heavy borrowing?

The most obvious feature of the sample languages is the one intuitively referred to as the 'grammar-lexicon split', which I have argued is better defined as a split between the INFL-language and the lexifier language of most of the potentially unbound core lexicon. Even Domari, with its significant lexical replacement through Arabic loans, does not match the Swadesh-list figures available for the languages dealt with here as MLs. One must however keep in mind that some of the MLs are better defined as registers (Para-Romani, Lekoudesch, Jenisch, and to some extent also Ma'á), and the preference of lexifier-language lexicon over INFL-language lexicon is a matter of discourse-strategic choice, with rather system-related constraints.

In addition, the sample MLs that have arisen in a situation of full bilingualism share the incorporation of lexifier-language deictics, interrogatives and pronominals, while those that serve as secret registers share a preference for lexifier-language copula and negators. In Michif and Media Lengua, lexifier-language nouns are accompanied by lexifier-language definite and indefinite articles, in Mednyj Aleut the nominal inflection (including converbs) is replicated from the lexifier language.

This – the sum of the 'core layer', the 'extended core layer', and 'accommodation layer' of the ML prototype – basically sums up the principal differences between the sample MLs and cases of attested heavy borrowing, such as Domari. Other properties found in the sample – the replication of some copula forms, negators, plural markers, numerals, and adpositions – are also found in Domari. Further borrowed features – conjunctions and discourse markers, and others not considered here – are not uncommon in conventional language contact situations.

MLs thus observe the constraint on the consistency of INFL-language – the language in which the predication is anchored and processed. They defy the normal constraint on the maintenance of a core balance between the source of the autonomous lexicon and the INFL-language. They further defy the normal constraint on harmony between the INFL-language and the source of deictics and pronominals. Finally, they allow for a violation of the constraint on harmony between INFL-language and nominal inflection (including case, agreement, definite and indefinite articles) if such a violation is functional in order to accommodate nominals from the lexifier-language on a wholesale basis. This behaviour of MLs contradicts the overwhelming tendencies even in cases of heavy, intense and prolonged contact; this is what makes MLs 'mixed'.

Which correlations can be observed within the sample? Apart from the typological profiles and potential clashes between the individual pairs of contributing languages, three factors appear to have some predictive relevance. The first is the direction of the process – INFL-language maintenance (with lexical re-orientation toward a lexifier language) versus INFL-language shift (selective replication from a lexifier language). In cases of INFL-language maintenance, the copula tends to pattern with the INFL-language. In cases of INFL-language shift, both the copula and deictics are replicated from the lexifier language.

The second factor is the availability of full bilingualism in the evolution of the languages. This is a pre-condition for the re-orientation or replication from the lexifier language of the possessive verb, of deictics, of definite articles and case inflection, but not for the substitution through lexifier-language elements of the copula, negators, or numerals (or personal pronouns, as demonstrated by secret languages not included in the sample).

The final factor is the function for which the languages are employed. In secret languages, the copula and negators pattern with the lexifier language but the possessive verb patterns with the INFL-language, while in plain languages the copula and negators may pattern with either source language, and often with both, while 'to have' is treated as a lexical verb and patterns with the source language for lexical verbs.

Acknowledgements

The findings presented here reflect work carried out as part of the research project on "Structural and functional constraints in cases of Stable Mixed Languages: Synchronic and diachronic aspects", funded by the Arts and Humanities Research Board under the Research Grant Scheme, and during a preparatory project funded by the Spinoza grant awarded to Pieter Muysken (University of Leiden). I acknowledge Anthony Grant's contribution to the compilation of part of the data from secondary sources. I am also grateful to Peter Bakker for helpful comments on the data and for inspiring discussions.

References

Auer, Peter
1999 From code-switching via language-mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3: 309–332.

Arends, Jacques, Pieter Muysken, and Norval Smith (eds.)
1995 *Pidgins and Creoles: An Introduction*. Amsterdam: John Benjamins.

Backus, Ad
1992 *Patterns of Language Mixing: A Study of Turkish–Dutch Bilingualism*. Wiesbaden: Harrassowitz.

Bakker, Peter
1994 Michif, the Cree–French mixed language of the Métis buffalo hunters in Canada. In Bakker and Mous (eds.), 13–33. Amsterdam: IFOTT.
1997 *"A Language of Our Own". The Genesis of Michif – The Mixed Cree–French Language of the Canadian Métis*. New York: Oxford University Press.
1998 Para-Romani languages and secret languages. In Matras (ed.), 69–96.
1994 Introduction. In Bakker and Mous (eds.), 1–11.

Bakker, Peter, and Maarten Mous (eds.)
1994 *Mixed Languages. 15 Case Studies in Language Intertwining*. Amsterdam: IFOTT [ordering address: aclc@mum.uva.nl].

Bakker, Peter, and Pieter Muysken
1995 Mixed languages and language intertwining. In Arends, Muysken, and Smith (eds.), 41–52.

Bakker, Peter, and Robert A. Papen
1997 Michif: A mixed language based on Cree and French. In Thomason (ed.), 295–363.

Bentahila, Abdelâli and Eirlys E. Davies
1995 Patterns of code-switching and patterns of language contact. *Lingua* 96: 75–93.

Berk-Seligson, Susan
1986 Linguistic constraints on intrasentential code-switching: A study of Spanish–Hebrew bilingualism. *Language in Society* 15: 313–348.

Boretzky, Norbert
1998 Der Romani-Wortschatz in den Romani-Misch-Dialekten. In Matras (ed.), 97–132.

Boretzky, Norbert, and Birgit Igla
1994 Romani mixed dialects. In Bakker and Mous (eds.), 35–68.

Campbell, Lyle
1993 On proposed universals of grammatical borrowing. In *Historical Linguistics 1989*, Henk Aertsen and Robert Jeffers (eds.), 91–110. Amsterdam: John Benjamins.

Dürr, Doris
1961 Ostschwäbische Händlerdörfer in Geschichte und Gegenwart. Volkskundliche Untersuchung zu den Dörfern Matzenbach und Unterdeufstetten im Kreis Crailsheim. Ph.D. dissertation, University of Tübingen.

Golovko, Evgenij V.
1994 Copper Island Aleut. In Bakker and Mous (eds.), 113–121.

Greenberg, Joseph H.
1999 Are there Mixed Languages? In *Essays in Poetics, Literary History and Linguistics: Presented to Viacheslav Vsevolodovich Ivanov on the Occasion of his Seventieth Birthday*, L. Fleishman, M. Gasparov, T. Nikolaeva, A. Ospovat, V. Toporov, A. Vigazin, R. Vroon, and A. Zaliznak (eds.), 626–633. Moscow: OGI.

Harris, Alice C., and Lyle Campbell
1995 *Historical Syntax in Cross-linguistic Perspective*. Cambridge: Cambridge University Press.

Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26: 210–231.

Igla, Birgit
1996 *Das Romani von Ajia Varvara*. Wiesbaden: Harrassowitz.

Jake, Janice L.
1994 Intrasentential code switching and pronouns: On the categorial status of functional elements. *Linguistics* 32: 271–298.

Joseph, Brian
1983 *The Synchrony and Diachrony of the Balkan Infinitive*. Cambridge: Cambridge University Press.

Klepsch, Alfred
1996 Das Lachoudische: Eine jiddische Sondersprache in Franken. In *Rotwelsch-Dialekte*, Klaus Siewert (ed.), 81–93. Wiesbaden: Harrassowitz.

Leigh, Kate
1998 Romani elements in present-day Caló. In Matras (ed.), 243–282.

Macalister, R. A. S.
1914 *The Language of the Nawar or Zutt, the Nomad Gypsy Smiths of Palestine*. London/Edinburgh: Edinburgh University Press.

Maschler, Yael
1998 On the transition from code-switching to a mixed code. In *Code-switching in Conversation. Languages, Interaction and Identity*, Peter Auer (ed.), 125–149. London: Routledge.

Masica, Colin P.
1976 *Defining a Linguistic Area: South Asia*. Chicago: Chicago University Press.

Matras, Yaron
1991 Zur Rekonstruktion des jüdischdeutschen Wortschatzes ehemaliger Judendörfer in Südwestdeutschland. *Zeitschrift für Dialektologie und Linguistik* 58, 268–293.
1996 Prozedurale Fusion: Grammatische Interferenzschichten im Romanes. *Sprachtypologie und Universalienforschung* 49 (1), 60–78.
1998 Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36, 281–331.
1999 The state of present-day Domari in Jerusalem, *Mediterranean Language Review* 11: 1–59
2000 Mixed Languages: A functional-communicative approach. *Bilingualism: Language and Cognition* 3, 79–99.
2002 *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
Matras, Yaron (ed.)
1998 *The Romani Element in Non-standard Speech*. Wiesbaden: Harrassowitz.
Moravcsik, Edith
1978 Universals of language contact. In *Universals of Human Language*, Joseph H. Greenberg (ed.), 94–122. Stanford: Stanford University Press.
Mous, Maarten
1994 Ma'á or Mbugu. In Bakker and Mous (eds.), 175–200.
Muysken, Pieter
1981 Halfway between Quechua and Spanish: The case for relexification. In *Historicity and Variation in Creole Studies*, A. Highfield and A. Valdman (eds.), 52–78. Ann Arbor: Karoma.
1994 Media Lengua. In Bakker and Mous (eds.), 207–211.
1997 Media Lengua. In Thomason (ed.), 365–426.
Myers-Scotton, Carol
1992 Codeswitching as a mechanism of deep borrowing, language shift, and language death. In *Language Death. Factual and Theoretical Explorations with Special Reference to East Africa*, Matthias Brenzinger (ed.), 31–58. Berlin: Mouton de Gruyter.
1993 *Duelling Languages: Grammatical Structure in Code-switching*. Oxford: Oxford University Press.
1998 A way to dusty death: The Matrix Language turnover hypothesis. In *Endangered Languages: Language Loss and Community Response*, L. A. Grenoble and L. J. Whaley (eds.), 289–316. Cambridge: Cambridge University Press.
Pfaff, Carol
1979 Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55: 291–318.

Poplack, Shana
1980 Sometimes I'll start a sentence in English y termino en español: Toward a typology of code-switching. *Linguistics* 18: 581–618.
Romaine, Suzanne
1989 *Bilingualism*. Oxford: Blackwell.
Ross, Malcolm
1996 Contact-induced change and the comparative method. In *The Comparative Method Reviewed*, Mark Durie and Malcolm Ross (eds.), 180–217. Oxford: Oxford University Press.
Rusakov, Alexander, and Olga Abramenko
1998 North Russian Romani dialect: Interference in case system. *Grazer Linguistische Studien* 50: 109–134.
Salmons, Joseph
1990 Bilingual discourse marking: Code switching, borrowing, and convergence in some German-American dialects. *Linguistics* 28, 453–480.
Silva-Corvalán, Carmen
1994 *Language Contact and Change*. Oxford: Clarendon Press.
Smith, Norval
1994 An annotated list of creoles, pidgins, and mixed languages. In Arends, Muysken and Smith (eds.), 331–374.
1998 Complex personal pronouns in the secret languages of nomadic castes. In *Mengelwerk voor Muysken*, Adriënne Bruyn and Jacques Arends (eds.), 44–49. Amsterdam: Institute for General Linguistics.
2000 Symbiotic mixed languages: A question of terminology. *Bilingualism Language and Cognition* 3.
Stolz, Christel, and Thomas Stolz
1996 Funktionswortentlehnung in Mesoamerika. Spanisch-amerindischer Sprachkontakt (Hispanoindiana II). *Sprachtypologie und Universalienforschung* 49 (1): 86–123.
Thomason, Sarah Grey
1995 Language mixture: Ordinary processes, extraordinary results. In *Spanish in Four Continents. Studies in Language Contact and Bilingualism*, Carmen Silva-Corvalán (ed.), 15–33. Washington D.C.: Georgetown University Press.
1997a Introduction. In Thomason (ed.), 1–7.
1997b Maa (Mbugu). In Thomason (ed.), 469–487.
1997c Mednyj Aleut. In Thomason (ed.), 449–468.
1997d A typology of contact languages. In *The Structure and Status of Pidgins and Creoles*, A. Spears, and D. Winford (eds.), 71–88. Amsterdam: John Benjamins.
2001 *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah Grey (ed.).
1997 *Contact Languages: A Wider Perspective*. Amsterdam: John Benjamins.
Thomason, Sarah Grey, and Terrence Kaufman
1988 *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
van Hout, Roeland, and Pieter Muysken
1994 Modeling lexical borrowability. *Language Variation and Change* 6: 39–62.
Weinreich, Uriel
1953 *Languages in Contact: Findings and Problems*. The Hague: Mouton.